ACL 2010 Handbook

The 48th Annual Meeting

of the Association for Computational Linguistics

July 11–16, 2010

Uppsala, Sweden

Contents

1 Welcome from the General Chair

Welcome back to Europe! After three years, the ACL crowd is meeting again in Europe, this time at the very north, to escape from the Central European heat it experienced in 2007.

This year, some significant changes can be found under the hood. The call for papers was formulated much more broadly than usual, and this idea brought up by the ACL membership and the Exec and then developed in detail by this year's program chairs, Sandra Carberry and Stephen Clark, really caught on – the number of submissions has been the highest of all times, forcing us to put some activities, such as the SRW, as the fifth track on Tuesday morning. The number of reviewers is hard to compute exactly – but a glimpse into their lists in this year's and previous years' proceedings reveals that we almost certainly set a new record here, too (thank you all!). Also, the proceedings have switched to electronic-only for all events, and adaptation of the START conference automation software has begun towards a fully automated workflow from submission to the production of the final proceedings in pdf format. It has been made possible thanks to Jing-Shin Chang's and Philipp Koehn's willingness to serve as Publication Chairs two years in a row in order to ensure a smooth transition from the semi-manual process employed in the past. However, there was one thing that overshadowed it all: the enthusiastic, meticulously precise and absolutely professional yet in every situation very polite approach of the local arrangements committee headed by Joakim Nivre. His efforts have made my job, as the General Chair, a piece of cake, limited essentially to watching the tons of emails exchanged between the local and other committees and to answering emails like "why wasn't I asked to be an invited speaker?" (obviously, from people no one would consider for this honor anyway).

Joakim has been helped by Beáta Megyesi, Rolf Carlson, Mats Dahllöf, Marco Kuhlmann, Mattias Nilsson, Markus Saers, Anna Sågvall Hein, Per Starbäck, Oscar Täckström, Jörg Tiedemann, Reut Tsarfaty and by the Akademikonferens team affiliated to Uppsala University headed by Ulla Conti; from her team, I would like to thank specifically Maria Carlson, Maria Bäckström and Johanna Thyselius Nilsson for taking care of the secretariat and website.

There are traditional ACL conference features as well – the workshops (with CoNLL-2010 as the big one), tutorials and the Student Research Workshop, the banquet (at the Uppsala Castle), the invited talks (albeit not-so-traditional this year, please come and see yourself), the Lifetime Achievement award, the business meeting and the closing session where the "conference torch" will be handed over to the Americas, as planned.

The Workshop Chairs (Pushpak Bhattacharyya and David Weir) had a hard time deciding which workshops to turn down, and tutorials had to be kept to a reason-

able number, too: quite an uneasy job for the Tutorials Chairs, Lluís Màrquez and Haifeng Wang. Demos have been selected by Sandra Kübler, and exhibitions handled by Jörg Tiedemann. Publicity has been the responsibility of Koenraad de Smedt and Beáta Megyesi, the local arrangements vice-chair. Students had again the opportunity to submit papers to the Student Research Workshop, organized by the SRW Chairs Nils Reiter, Seniz Demir, and Jan Raab, helped by their Faculty Advisor Tomek Strzalkowski, who also handled the application for the usual NSF grant supporting the SRW. Markéta Lopatková, the other Faculty Advisor, then centrally handled the student travel grants. Mentoring was the responsibility of Björn Gambäck and Diana McCarthy. Talking about money and the budget, the sponsoring committee has been quite successful this year by securing grants both locally and internationally: Mats Wirén, Hercules Dalianis, Christy Doran, Srinivas Bangalore, Frédérique Segond, and Stephen Pulman assembled an impressive lineup of sponsors. Thanks to them, all of you can benefit from low registration fees, subsidized banquet, the conference bag, and student scholarships and prizes.

No thank-you would be complete without mentioning Priscilla Rasmussen – her experience, insight, and ability to predict the numbers and other things was extremely helpful, to say the least. The ACL treasurer, Graeme Hirst, has helped to reassure us whenever there was doubt or an open budgetary question. And Steven Bird, who chaired the coordinating committee (a subcommittee of the ACL and EACL executive boards) which selected the conference venue and appointed the general chair and program chairs, has been with us throughout almost two years of preparations, helping to make sure we all (read and) follow the Conference organization handbook, and address all possible problems.

Finally, a conference without papers (and you as participants, of course) would not happen at all. Thank you for working hard, for submitting solid work and for preparing interesting talks and posters!

Enjoy the conference.

Jan Hajič
ACL 2010 General Chair
July 2010

2 About the Conference in Uppsala

For the first time in its nearly 50-year history, the flagship conference of the Association of Computational Linguistics (ACL) is being held in Scandinavia, in the city of Uppsala, Sweden. ACL 2010 will cover a broad spectrum of areas related to natural language and computation, and you are welcome to participate in the conference to discuss the latest research findings during the Main Conference, the Student Research Workshop, and the System Demonstrations. As is customary, ACL 2010 is preceded by one day of Tutorials and followed by two days of Workshops. Collocated with ACL in Uppsala is also the Fourteenth Conference on Computational Natural Language Learning (CoNLL).

With this exciting line-up of events, we welcome members of many research communities to the city of Uppsala. First mentioned in the Beowulf saga dating back to the 6th century, Uppsala has for long periods been the political, religious and academic center of Sweden. It is the seat of the archbishop of the Church of Sweden since 1164 and the seat of the oldest university in Scandinavia founded in 1477. Today Uppsala is Sweden's fourth largest city with a population of 200 000 inhabitants and has retained its small-town charm while offering a big city's selection of shops, restaurants and other entertainment. It is a city with unique cultural treasures and historical attractions, including the largest cathedral in Scandinavia, a castle from the 16th century, which will house the ACL 2010 Banquet, the Linnaeus garden, and a unique anatomical theater from the 17th century – all within easy walking distance in the city center.

The venue for ACL 2010 is the Uppsala University Main Building (Venue A) and the nearby Center for Economic Studies – Ekonomikum (Venue B). The Main Building, built in roman renaissance style during the second half of the 19th century, will be the venue for tutorials and most events associated with the main conference, including all plenary sessions, while the modern Ekonomikum building will be used for some parallel sessions during the main conference and a few workshops. The two buildings are within five minutes walking distance of each other, and breaks have been inserted at appropriate places in the program to allow participants to move from one building to the other and minimize the inconvenience caused by the split venue. We wish you all a successful and enjoyable conference in Uppsala!

Joakim Nivre
ACL 2010 Local Arrangements Chair
July 2010

3 Organization

3.1 ACL

The Association for Computational Linguistics is *the* international scientific and professional society for people working on problems involving natural language and computation. Membership includes (among other things) reduced registration at most ACL-sponsored conferences, discounts on publications of participating publishers and ACL and related publication back issues, announcements of ACL and related conferences, workshops, and journal calls of interest to the community, and participation in ACL Special Interest Groups.

The ACL journal, C*omputational Linguistics*, continues to be the primary forum for research on computational linguistics and natural language processing. It is now published electronically and all its papers appear in the ACL Anthology which provides open access also to all papers from ACL-sponsored conferences and workshops.

An annual meeting is held each summer in locations where significant computational linguistics research is carried out.

3.2 CoNLL

The Conference on Computational Natural Language Learning (CoNLL) is the yearly international conference on natural language learning organized by SIGNLL (the ACL Special Interest Group on Natural Language Learning). The conference typically has a special topic of interest which this year is on grammar induction. CoNLL is accompanied every year by a shared task intended to promote natural language processing applications and evaluate them in a standard setting. The 2010 shared task is on learning to detect hedges and their scope in natural language texts.

3.3 Committees

Organizing Committee

General Conference Chair
Jan Hajič (Charles University, Czech Republic)

Program Chairs
Sandra Carberry (University of Delaware, USA)
Stephen Clark (University of Cambridge, UK)

Local Arrangements Chair
Joakim Nivre (Uppsala University, Sweden)

Workshop Chairs
Pushpak Bhattacharyya (Indian Institute of Technology, India)
David Weir (University of Sussex, UK)

Tutorial Chairs
Lluís Màrquez (Technical University of Catalonia, Spain)
Haifeng Wang (Baidu.com Inc., China)

System Demonstration Chair
Sandra Kübler (Indiana University, USA)

Student Research Workshop Committee
Seniz Demir (University of Delaware, USA)
Jan Raab (Charles University, Czech Republic)
Nils Reiter (Heidelberg University, Germany)

Student Research Workshop Faculty Advisors
Marketa Lopatkova (Charles University, Czech Republic)
Tomek Strzalkowski (State University of New York, USA)

Publications Chairs
Jing-Shin Chang (National Chi-Nan University, Taiwan)
Philipp Koehn (University of Edinburgh, UK)

Mentoring Service Chairs
Björn Gambäck (SICS, Sweden and NTNU, Norway)
Diana McCarthy (Lexical Computing Ltd., UK)

Sponsorship Chairs
Stephen Pulman (University of Oxford, UK)
Frédérique Segond (Xerox Research Centre Europe, France)
Srinivas Bangalore (AT&T Research, USA)
Christy Doran (MITRE, USA)
Hercules Dalianis (Stockholm University/KTH, Sweden)
Mats Wirén (Stockholm University, Sweden)

Publicity Chairs
Koenraad de Smedt (University of Bergen, Norway)
Beáta Megyesi (Uppsala University, Sweden)

Exhibition Chair
Jörg Tiedemann (Uppsala University, Sweden)

Local Arrangements Committee
Joakim Nivre (Uppsala University, Sweden)
Rolf Carlson (KTH, Sweden)
Mats Dahllöf (Uppsala University, Sweden)
Marco Kuhlmann (Uppsala University, Sweden)
Beáta Megyesi (Uppsala University, Sweden)
Mattias Nilsson (Uppsala University, Sweden)
Markus Saers (Uppsala University, Sweden)
Anna Sågvall Hein (Uppsala University, Sweden)
Per Starbäck (Uppsala University, Sweden)
Jörg Tiedemann (Uppsala University, Sweden)
Reut Tsarfaty (Uppsala University, Sweden)
Oscar Täckström (Uppsala University, Sweden)

Secretariat, Webmaster
Academic Conferences (Uppsala University, Sweden):
Ulla Conti
Maria Carlson
Maria Bäckström
Johanna Thyselius Nilsson

Registration
Priscilla Rasmussen (ACL)

Program Committee

For ACL

Program Co-Chairs
Sandra Carberry (University of Delaware, USA)
Stephen Clark (University of Cambridge, UK)

Area Chairs
Tim Baldwin (University of Melbourne, Australia)
Phil Blunsom (University of Oxford, UK)
Kalina Bontcheva (University of Sheffield, UK)
Johan Bos (University of Rome – La Sapienza, Italy)
Claire Cardie (Cornell University, USA)
Walter Daelemans (University of Antwerp, Belgium)
Robert Gaizauskas (University of Sheffield, UK)
Keith Hall (Google Research – Zurich, Switzerland)
Julia Hirschberg (Columbia University, USA)
Nancy Ide (Vassar College, USA)
Michael Johnston (AT&T Labs, USA)
Roger Levy (University of California – San Diego, USA)
Hang Li (Microsoft Research Asia, China)
Chin-Yew Lin (Microsoft Research Asia, China)
Yusuke Miyao (University of Tokyo, Japan)
Roberto Navigli (University of Rome – La Sapienza, Italy)
Ani Nenkova (University of Pennsylvania, USA)
Jon Oberlander (University of Edinburgh, UK)
Chris Quirk (Microsoft Research, USA)
Stuart M. Shieber (Harvard University, USA)
Khalil Sima'an (University of Amsterdam, The Netherlands)
Richard Sproat (Oregon Health and Science University, USA)
Matthew Stone (Rutgers University, USA)
Jun'ichi Tsujii (University of Tokyo, Japan, and University of Manchester, UK)
Bonnie Webber (University of Edinburgh, UK)
Theresa Wilson (University of Edinburgh, UK)
ChengXiang Zhai (University of Illinois at Urbana–Champaign, USA)

For CoNLL

Conference Chairs
Anoop Sarkar (Simon Fraser University, Canada)
Mirella Lapata (University of Edinburgh, UK)

Local Arrangements Committee

Local Arrangements Chair
Joakim Nivre (Uppsala University, Sweden)

Vice Chair, Conference Handbook, Newsletters, Publicity
Beáta Megyesi (Uppsala University, Sweden)

Conference Handbook
Mats Dahllöf (Uppsala University, Sweden)

Posters, Demos
Marco Kuhlmann (Uppsala University, Sweden)

Student Volunteer Programme
Mattias Nilsson (Uppsala University, Sweden)

Banquet, Conference Handbook
Markus Saers (Uppsala University, Sweden)

Wireless Internet, Technical Support
Per Starbäck (Uppsala University, Sweden)

Exhibits, Sponsorship
Jörg Tiedemann (Uppsala University, Sweden)

Workshops and CoNLL
Reut Tsarfaty (Uppsala University, Sweden)

Webmaster, Secretariat, Social Events
Academic Conferences (Uppsala University, Sweden)

Graphical Design
Södra Tornet (Sweden)

4 Information

4.1 Conference Venue

The conference takes place at Uppsala University in a genuine university environment, dating back as far as 1477. All technical sessions are held either in the Uppsala University Main Building (Venue A) or the Center for Economic Studies (Venue B). Venue A is located in the University Park, at the corner of Övre Slottsgatan/S:t Olofsgatan marked as No. 1 on the city map. Venue A is 10–15 minutes walking distance from the central railway station. Venue B, marked as No. 2 on the city map, is located at Kyrkogårdsgatan 10 which is within 5 minutes walking distance from Venue A. Venue and city maps with important locations marked can be found at the end of the handbook. You may check the assigned presentation room and the session timetable at the conference website for updates.

4.2 Instructions for Presenters

Oral Presentations

The following instructions for presenters are applicable for all oral sessions in the ACL Main Conference, CoNLL and Workshops.

Equipment

Each presentation room is equipped with a laptop computer, a data projector, a microphone (for large rooms), a lectern, and a pointing device. You are strongly recommended to use the laptops provided by the conference.

Identical laptops with the same specifications are also available in the Speaker Ready Room (Room 2) during the main conference. You can check if your slides can be displayed properly in the Speaker Ready Room.

The laptops are equipped with:

- Windows XP SP3
- Internet connection, USB port, DVD player
- Microsoft Office 2007
- Adobe Reader, Flash Player, Media Players (Microsoft/Real/QuickTime)
- Anti-Virus software

You are advised to check if your PowerPoint slides can be displayed properly using PowerPoint Viewer 2007. The computers used for presentations will have wired internet. WiFi is also available at the conference venue, however, the bandwidth is only enough for web browsing and email, not for video/audio streaming.

Presentation

Your slides should be uploaded to the laptop in your session room. This should be done half an hour prior to the start of the first morning session (for morning presentations) or half an hour prior to the end of the lunch recess (for afternoon presentations). Please arrive at your session at least 15 minutes prior to the start of your session; you should introduce yourself to the session chair and ask if there are any last-minute instructions.

Long talks are allotted 20 minutes for presentation and 5 minutes for questions from the audience. Please ensure that your presentation does not exceed 20 minutes in length. Your presentation should highlight the problem(s) addressed by your research, describe the approach/methodology used to address these problems, discuss evaluation of your results, and compare your work to other research.

Short talks are allotted 9 minutes for presentation; questions will be held until the associated poster session. Please ensure that your presentation does not exceed 9 minutes in length. Presentations that exceed 9 minutes will be stopped by the session chair in order to allow the other presenters to have their full allotted time. Your presentation should highlight the key points of your research and its novel contributions, and encourage the audience to attend your poster presentation for further details about your work.

The allocated presentation time for workshops and CoNLL may differ. Please check the conference web site for the exact time allocation for your presentation.

Poster and Demo Presentations

These presentation instructions are applicable for posters and demonstrations in the ACL Main Conference, Student Research Workshop, CoNLL and Workshops.

Equipment

For posters, we will provide display easels measuring 100cm in width and 140cm in height, with a usable board area of 95cm x 135cm. This size is good for a standard A0 poster in the portrait orientation. The poster easels are double-sided with one poster on each side. Pins for mounting will be provided. However, no tables will be available except for Software Demonstrations.

For software demonstrations, we will provide a table with a chair, an electric outlet (220/230 V), a poster easel (usable board area 95cm x 135cm), and a 10 Mbit/s Ethernet connection (via a standard 8P8C/RJ45 outlet).

Presentation

A paper presented as a poster offers a unique opportunity to present research work in a way customized to individuals or a small group of people. It is more interactive than an oral presentation. Therefore, the work can be presented, in certain respects, more effectively to a small but well-targeted audience. Remember people attracted by a poster are so interested in the work that they are willing to invest anywhere from 5 to 10 minutes of their time. That is a big chunk out of their time at a poster session!

To attract the audience who would be interested in your work, the poster should have a title in large font which is clearly visible to even passers-by. Its contents should also be in fonts large enough to be readable from 1 to 2 meters away. Instead of

constructing your poster as an enlarged summary of an oral presentation, you should take advantage of the flexibility that a poster offers with respect to organization. For example, you might want to place a system diagram in the center, surrounded by descriptions and performance tables of its individual components. Or you might want to place an example in the center, with arrows to the problems it illustrates and the methodologies used to address these problems. The best posters will take advantage of this flexibility.

"A picture is worth a thousand words". Try to choose visual aids like figures, diagrams, cartoons, colors, even lines over texts on your poster to show the research idea and the logical flow of the contents. Thus after attracting attendees with an enticing title, the poster can be self-explaining so that people can understand it and quickly find out whether they have more questions to ask. If they do, they can have a short discussion with you to get the most out of your poster presentation. In addition, some people are more verbal than visual. They prefer to listen instead of read, even when the visualization is great. So, prepare "mini-talks" as short as 30 seconds, and some as long as 5 minutes. Kindly ask people (who might appear to be reading the poster slowly) whether they would like a brief introduction from you. You will need to adapt to your audience. Senior researchers in your area of expertise probably need only a few key points explained, while more general information would help those not so familiar with your task. Please try to interact with everyone who seems interested in your work, rather than have long intricate conversations with a few. If someone wants to discuss your work in extensive detail, this is a great opportunity to arrange an individual meeting later in the conference.

Occasionally, people prepare printouts to complement their posters. If you expect such printouts to be helpful, please prepare them.

Please avoid leaving your poster without a presenter, since then it will attract less attention than it deserves.

4.3 Awards

Best Paper Awards

In the ACL tradition, ACL 2010 has several Best Paper awards: best long paper, best short paper, and best long paper authored by a student, sponsored by IBM. The best long paper will receive its own plenary session at the end of the conference, and the recipients of the prizes will each receive a certificate and cash award. The Best Paper awards have been decided by a Best Paper committee, consisting of some of the members of the Program Committee and additional members drawn from the ACL community.

Best long paper

Beyond NomBank: A Study of Implicit Arguments for Nominal Predicates
by Matthew Gerber and Joyce Chai.
Presented: Wednesday July 14, 2010 17:50–18:15, Venue A, Aula.

Best short paper

SVD and Clustering for Unsupervised POS Tagging
by Michael Lamar, Yariv Maron, Mark Johnson and Elie Bienenstock.
Presented: Tuesday July 13, 2010 12:15–12:25, Venue A, Hall X.

IBM Best student paper

Extracting Social Networks from Literary Fiction
by David Elson, Nicholas Dames and Kathleen McKeown.
Presented: Monday July 12, 2010, 11:20–11:45, Venue B, Hall 4.

Lifetime Achievement Award

The ACL Lifetime Achievement Award (LTA) was instituted on the occasion of the Association's 40^{th} anniversary meeting. The award is presented for scientific achievement, of both theoretical and applied nature, in the field of Computational Linguistics. Currently, an ACL committee nominates and selects at most one award recipient annually, considering the originality, depth, breadth, and impact of the entire body of the nominee's work in the field. The award is a crystal trophy and the recipient is invited to give a 45-minute speech on his or her view of the development of Computational Linguistics at the annual meeting of the association. As of 2004, the speech has been subsequently published in the Association's journal, *Computational Linguistics*. The speech is introduced by the announcement of the award winner, whose identity is not made public until that time.

Previous winners of the distinguished award have been: Aravind Joshi (2002), Makoto Nagao (2003), Karen Spärck Jones (2004), Martin Kay (2005), Eva Hajičová (2006), Lauri Karttunen (2007), Yorick Wilks (2008) and Fred Jelinek (2009).

4.4 Practical Information

Registration and Information

The registration and main information desk/secretariat is located at the Uppsala University main building (Venue A). Registration will open on Sunday, July 11.

Opening hours for the secretariat and registration desk:

- Sunday July 11: 7:30–18:00
- Monday July 12: 7:00–17:30
- Tuesday July 13: 7:30–17:30
- Wednesday July 14: 7:30–17:30
- Thursday July 15: 7:30–17:30
- Friday July 16: 7:30–14:30

The local information desk will be open:

- Sunday July 11: 9:00–16:00
- Monday July 12: 8:00–16:00
- Tuesday July 13: 8:00–16:00
- Wednesday July 14: 8:00–16:00

Phone: +46 (0) 730 23 84 32 (available only during opening hours of the local information desk).

Members of the local arrangements committee and the local student volunteers can be identified by a big red dot on their name badge. Student volunteers will wear red T-shirts with the text "ACL 2010 CREW".

Name Badge

Your name badge is your admission to the scientific sessions as well as to the coffee and lunches. You should wear it at all times at the conference venue.

Coffee/Tea

Coffee and tea will be served during the breaks according to the program in Venue A and Venue B. Note that during the first morning breaks (between 10:00 and 10:30) of the main conference, coffee/tea is served only at Venue A. Please see the program for time and place. You need your name badge as a ticket.

Lunches

Lunch is included in the registration fee for all main conference participants during the main conference. The lunch will be served in the foyer of Venue A. You need your name badge as a ticket. Note that lunch is not included during the tutorials and workshops.

Exhibitors

The exhibitors can be found in the foyer of Venue A during the main conference.

Internet Access

Wireless Internet access will be available for conference participants in the conference area. You can login through Eduroam, or use a personal login which you will receive with your name badge when registering on site.

Eduroam

Eduroam is an encrypted roaming access service developed for the research and education community, available in most of Europe and some other parts of the world. You will get your account from your local network administrator beforehand if you don't already have it. You can read about Eduroam at `http://www.eduroam.org`.

UpUnet-S

If you don't have Eduroam you can use a personal temporary guest account here for the network "UpUnet-S" instead. You get login credentials for that when registering. Notice: Traffic is not encrypted.

4.5 Social Events

Welcome Reception – July 11, 18:00–21:00

The Welcome Reception is free for all ACL 2010 registered participants. Drinks and light snacks will be served. Please join us in the foyer of the Uppsala University Main Building (Venue A) to meet old and new friends!

Main Conference Lunch – July 12–14

A light lunch is included in the main conference registration and will be served in the foyer of the Uppsala University Main Building (Venue A) on all three days of the main conference. On Monday and Tuesday, lunch is combined with the main conference poster sessions from 13:15 to 15:00. On Wednesday, lunch is served from 13:00 to 14:30, partly overlapping with the ACL Business Meeting (12:20–13:20).

Student Club – July 12–14

Because lunch is included in the registration fee for all main conference participants, there will be no traditional Student Lunch at ACL 2010. But during the main conference, Room II in the Main Building (Venue A) will be reserved as a Student Club, where students attending ACL can meet their peers for birds-of-a-feather meetings, impromptu demos, or just a social chat.

Banquet – July 13, 19:00–01:00

The ACL 2010 Banquet will be held at Uppsala Castle (Uppsala slott), in the Hall of State (Rikssalen). The construction of the castle was started in 1549 by king Gustav I, but it has been remodeled and expanded on several occasions, in particular after a great fire in 1702. Throughout much of its early history, Uppsala Castle played a major role in the history of Sweden, being the seat of the Swedish parliament and the scene of many important political events. Today the castle is the residence of the County Governor of Uppsala County, and the Hall of State is the supreme banquet hall in Uppsala. The ACL 2010 Banquet will feature a three-course dinner, followed by dancing to live music according to the established ACL tradition. Uppsala Castle is marked as No. 14 on the city map. Pre-registration is required.

4.6 Local Information

Emergency Phone Numbers

You should call 112 if anything happens which means that an ambulance, the police or the fire brigade need to be called out. 112 is a special emergency number you can call wherever you are from a fixed or a mobile telephone.

International Calls

Dial 00 + country code + area code + phone number. For example, to call the US, dial 00+1, Germany 00+49.

Medical Services

Uppsala University hospital, Akademiska sjukhuset, is located in central Uppsala and marked with No. 17 on the city map. Telephone: +46 18 611 0000. The emergency room is called *Akuten* in Swedish.

Pharmacy

There are several pharmacies in Uppsala. Look for a sign *Apotek*. One is marked on the city map with No. 16. Grocery stores and other retail outlets might also sell certain non-prescription drugs to customers over the age of 18.

Money Exhange, Currency

Swedish Krona (SEK) is the official currency in Sweden. An exchange office is available next to the tourist office (Forex), marked with No. 15 on the city map. There are plenty of cash dispensers in Uppsala. Major international credit cards are accepted in most hotels, shops and restaurants.

Shopping

Most stores in Uppsala are open 10.00-19.00 on weekdays and 10.00-17.00 on Saturdays. Some stores are open on Sundays as well. Grocery stores usually have longer opening hours.

Eating

There are plenty of restaurants in Uppsala. You find information about some of them which we recommend in your conference bag.

Electricity

In Sweden the Europlug (Type C & F) is used for electricity with two round prongs and the electrical voltage is 220/230V. For more information, see http://en.wikipedia.org/wiki/AC_power_plugs_and_sockets.

Smoking

Smoking is not allowed in the conference venues, nor in any public indoor establishments such as restaurants and bars.

Transportation to Stockholm Arlanda International Airport (ARN)

Taxi

You can prebook a taxi at (+46 18) 100 000 as well as at (+46 18) 123 456, or at www.uppsalataxi.se. The price to get to Stockholm Arlanda International Airport is about 460 SEK. The journey takes approximately 30 minutes.

Bus

Bus 801 runs between Uppsala Central Station and Arlanda twice/hour from about 4 am until midnight, and once/hour from midnight until 4 am. The journey takes about 45 minutes and costs 100 SEK. You can buy your ticket from the driver paying in SEK.

Train

Trains leave from Uppsala Central Station to Stockholm Arlanda International Airport 1–3 times/hour from 5:00 until 23:00. The journey takes 15–20 minutes and costs 100–140 SEK if purchased in advance at Uppsala Central Station.

If you need help to get to another airport, please ask the local information desk in the main university building (Venue A).

Platinum Sponsors

Riksbankens Jubileumsfond

Riksbankens Jubileumsfond is an independent foundation with the goal of promoting and supporting research in the Humanities and Social Sciences.

Vetenskapsrådet

The Swedish Research Council (Vetenskapsrådet) is a government agency that provides funding for basic research of the highest scientific quality in all disciplinary domains. Besides research funding, the agency works with strategy, analysis, and research communication.

UPPSALA
UNIVERSITET

Uppsala University is an internationally prominent research university that is a world leader in many areas. The University offers both breadth and depth in its subject areas. With a tradition of research and education stretching back over 500 years, Uppsala University is constantly seeking new approaches. The University's tradition of renewal is one of its strengths.

Gold Sponsors

GSLT
Swedish National Graduate School of Language Technology
www.gslt.hum.gu.se

GSLT is coordinated by the University of Gothenburg and is a collaboration with the following universities: Borås, Chalmers, KTH, Linköping, Linnaeus, Lund, Skövde, Stockholm and Uppsala.

The school integrates research on speech and language and provides a sound basis in both theoretical foundations and applications oriented research. Courses are taught in English to an international student body. Currently 24 PhDs have been completed.

Silver Sponsors

Bronze Sponsors

Other Sponsors

NICE Systems
Supporter

Language Weaver
Conference Bag Sponsor

Xerox Research Centre Europe
Overseas Student Fellowship Sponsor

Swedish Institute of Computer Science (SICS)
Local Student Fellowship Sponsor

National Science Foundation (NSF)
Sponsor of the Student Research Workshop

IBM Research
Sponsor of Best Student Paper Award

Exhibitors

Acapela Group
http://www.acapela-group.com/

Morgan & Claypool Publishers
http://www.morganclaypool.com/

Swedish National Graduate School of Language Technology (GSLT)
http://www.gslt.hum.gu.se/

textkernel
http://www.textkernel.nl/

5 Program at a Glance

Day	Activities	Time	Venues
Sunday, July 11	Registration	7:30–18:00	A
	Tutorials	9:00–17:30	A
	Welcome Reception	18:00–21:00	A
Monday, July 12	Registration	7:00–17:30	A
	Main Conference, Day 1	8:45-18:00	A & B
Tuesday, July 13	Registration	7:30–17:30	A
	Main Conference, Day 2	9:00–17:35	A & B
	Student Research Workshop	10:30–13:15	A
	Demo Session	15:00–17:35	A
	Banquet at Uppsala Castle	19:00–01:00	
Wednesday, July 14	Registration	7:30–17:30	A
	Main Conference, Day 3	9:00–18:30	A & B
	ACL Business Meeting	12:20–13:20	A
Thursday, July 15	Registration	7:30–17:30	A
	CoNLL, Day 1	9:00–18:00	A
	Workshops, Day 1		
	WS 1, 2, 3, 4, 5, 7, 11 and 12	9:00–17:30	A & B
Friday, July 16	Registration	7:30–14:30	A
	CoNLL, Day 2	9:15–17:45	A
	Workshops, Day 2		
	WS 1, 2, 3, 6, 8, 9, 10 and 13	9:00–17:30	A & B

Start and end times for workshops may vary.

Venue A: Uppsala University Main Building ("Universitetshuset").

Venue B: Center for Economic Studies ("Ekonomikum").

There is a 5 minutes walking distance between Venue A and Venue B.

6 Tutorials, July 11

	Venue A, Hall X	Venue A, Hall IX	Venue A, Hall IV
9:00–10:30	**T5: Tree-based and Forest-based Translation** Yang Liu and Liang Huang	**T3: Wide-Coverage NLP with Linguistically Expressive Grammars** Josef van Genabith, Julia Hockenmaier and Yusuke Miyao	**T6: Discourse Structure: Theory, Practice and Use** Bonnie Webber, Markus Egg and Valia Kordoni
10:30–11:00	*Break*		
11:00-12:30	**T5, continued**	**T3, continued**	**T6, continued**
12:30-14:00	*Lunch*		
14:00-15:30	**T2: From Structured Prediction to Inverse Reinforcement Learning** Hal Daumé III	**T4: Semantic Parsing: The Task, the State of the Art and the Future** Rohit J. Kate and Yuk Wah Wong	**T1: Annotation** Eduard Hovy
15:30-16:00	*Break*		
16:00-17:30	**T2, continued**	**T4, continued**	**T1, continued**

T1: Annotation

July 11, 14:00–17:30, Venue A, Hall IV

Eduard Hovy

As researchers seek to apply their machine learning algorithms to new problems, corpus annotation is increasingly gaining importance in the NLP community. But since the community currently has no general paradigm, no textbook that covers all the issues (though Wilcock's book published in 2009 covers some basic ones very well), and no accepted standards, setting up and performing small-, medium-, and large-scale annotation projects remain somewhat of an art.

This tutorial is intended to provide the attendee with an in-depth look at the procedures, issues, and problems in corpus annotation, and highlights the pitfalls that the annotation manager should avoid. The tutorial first discusses why annotation is becoming increasingly relevant for NLP and how it fits into the generic NLP method-

ology of train-evaluate-apply. It then reviews currently available resources, services, and frameworks that support someone wishing to start an annotation project easily. This includes the QDAP annotation center, Amazon's Mechanical Turk, annotation facilities in GATE, and other resources such as UIMA. It then discusses the seven major open issues at the heart of annotation for which there are as yet no standard and fully satisfactory answers or methods. Each issue is described in detail and current practice is shown. The seven issues are: 1. How does one decide what specific phenomena to annotate? How does one adequately capture the theory behind the phenomenon/a and express it in simple annotation instructions? 2. How does one obtain a balanced corpus to annotate, and when is a corpus balanced (and representative)? 3. When hiring annotators, what characteristics are important? How does one ensure that they are adequately (but not over- or under-) trained? 4. How does one establish a simple, fast, and trustworthy annotation procedure? How and when does one apply measures to ensure that the procedure remains on track? How and where can active learning help? 5. What interface(s) are best for each type of problem, and what should one know to avoid? How can one ensure that the interfaces do not influence the annotation results? 6. How does one evaluate the results? What are the appropriate agreement measures? At which cutoff points should one redesign or redo the annotations? 7. How should one formulate and store the results? When, and to whom, should one release the corpus? How should one report the annotation effort and results for best impact?

Eduard Hovy is the director of the Natural Language Group at USC/ISI. His research focuses on questions in information extraction, automated text summarization, the semi-automated construction of large lexicons and ontologies, machine translation, question answering, and digital government. Much of this work has required annotation. Together with colleagues, students, and visitors, he has had annotation projects in biomedical information extraction, coreference, word-sense annotation, ontology creation, noun-noun relations, and discourse structure. The smallest of these projects (discourse structure) involved three annotators over a period of three months, and the largest (OntoNotes noun senses) involved more than 25 annotators over several years. Some of these projects used the CAT annotation interface developed at UPitt, others involved home-grown interfaces, and some of them involved Amazon's Mechanical Turk.

T2: From Structured Prediction to Inverse Reinforcement Learning

July 11, 14:00–17:30, Venue A, Hall X

Hal Daume III

Machine learning is all about making predictions; language is full of complex rich structure. Structured prediction marries these two. However, structured prediction isn't always enough: sometimes the world throws even more complex data at us, and we need reinforcement learning techniques. This tutorial is all about the *how* and the *why* of structured prediction and inverse reinforcement learning (aka inverse optimal control): participants should walk away comfortable that they could implement many structured prediction and IRL algorithms, and have a sense of which ones might work

for which problems.

The first half of the tutorial will cover the "basics" of structured prediction the structured perceptron and Magerman's incremental parsing algorithm. It will then build up to more advanced algorithms that are shockingly reminiscent of these simple approaches: maximum margin techniques and search-based structured prediction.

The second half of the tutorial will ask the question: what happens when our standard assumptions about our data are violated? This is what leads us into the world of reinforcement learning (the basics of which we'll cover) and then to inverse reinforcement learning and inverse optimal control.

Throughout the tutorial, we will see examples ranging from simple (part of speech tagging, named entity recognition, etc.) through complex (parsing, machine translation).

The tutorial does not assume attendees know anything about structured prediction or reinforcement learning (though it will hopefully be interesting even to those who know some!), but *does* assume some knowledge of simple machine learning (eg., binary classification).

Hal Daume III is an assistant professor in the School of Computing at the University of Utah. His primary research interests are in understanding how to get human knowledge into a machine learning system in the most efficient way possible. In practice, he works primarily in the areas of Bayesian learning (particularly non-parametric methods), structured prediction and domain adaptation (with a focus on problems in language and biology). He associates himself most with conferences like ACL, ICML, NIPS and EMNLP. He earned his PhD at the University of Southern California with a thesis on structured prediction for language (his advisor was Daniel Marcu). He spent the summer of 2003 working with Eric Brill in the machine learning and applied statistics group at Microsoft Research. Prior to that, he studied math (mostly logic) at Carnegie Mellon University.

T3: Wide-Coverage NLP with Linguistically Expressive Grammars

July 11, 09:00–12:30, Venue A, Hall IX

Julia Hockenmaier, Yusuke Miyao and Josef van Genabith

In recent years, there has been a lot of research on wide-coverage statistical natural language processing with linguistically expressive grammars such as Combinatory Categorial Grammars (CCG), Head-driven Phrase-Structure Grammars (HPSG), Lexical-Functional Grammars (LFG) and Tree-Adjoining Grammars (TAG). But although many young researchers in natural language processing are very well trained in machine learning and statistical methods, they often lack the necessary background to understand the linguistic motivation behind these formalisms. Furthermore, in many linguistics departments, syntax is still taught from a purely Chomskian perspective. Additionally, research on these formalisms often takes place within tightly-knit, formalism-specific subcommunities. It is therefore often difficult for outsiders as well as experts to grasp the commonalities of and differences between these formalisms.

This tutorial overviews basic ideas of TAG/CCG/LFG/HPSG, and provides at-

tendees with a comparison of these formalisms from a linguistic and computational point of view. We start from stating the motivation behind using these expressive grammar formalisms for NLP, contrasting them with shallow formalisms like context-free grammars. We introduce a common set of examples illustrating various linguistic constructions that elude context-free grammars, and reuse them when introducing each formalism: bounded and unbounded non-local dependencies that arise through extraction and coordination, scrambling, mappings to meaning representations, etc. In the second half of the tutorial, we explain two key technologies for wide-coverage NLP with these grammar formalisms: grammar acquisition and parsing models. Finally, we show NLP applications where these expressive grammar formalisms provide additional benefits.

Julia Hockenmaier is assistant professor in the Department of Computer Science at the University of Illinois, Urbana-Champaign. She has been working on translating the English Penn Treebank and the German Tiger corpus to CCG, and developed one of the first statistical parsers for CCG.

Yusuke Miyao is an associate professor in National Institute of Informatics, Japan. He has been engaged in the research of wide-coverage HPSG parsing, specifically focusing on statistical models for parse disambiguation, and the treebank-based development of wide-coverage grammars. He has also been working on the applications of the HPSG parser, including biomedical IE/IR and wide-coverage logical form construction.

Josef van Genabith is an associate professor in the School of Computing at Dublin City University and the director of the Centre for Next Generation Localisation (CNGL). He has been working on treebank-based acquisition of wide-coverage LFG resources (for English, German, Spanish, French, Chinese, Arabic and Japanese) and data-driven parsing and generation models for these resources.

T4: Semantic Parsing: The Task, the State of the Art and the Future

July 11, 14:00–17:30, Venue A, Hall IX

Rohit J. Kate and Yuk Wah Wong

Semantic parsing is the task of mapping natural language sentences into complete formal meaning representations which a computer can execute for some domain-specific application. This is a challenging task and is critical for developing computing systems that can understand and process natural language input, for example, a computing system that answers natural language queries about a database, or a robot that takes commands in natural language. While the importance of semantic parsing was realized a long time ago, it is only in the past few years that the state-of-the-art in semantic parsing has been significantly advanced with more accurate and robust semantic parser learners that use a variety of statistical learning methods. Semantic parsers have also been extended to work beyond a single sentence, for example, to use discourse contexts and to learn domain-specific language from perceptual contexts. Some of the future research directions of semantic parsing with potentially large impacts include mapping entire natural language documents into machine processable form to enable automated reasoning about them and to convert natural language web pages

into machine processable representations for the Semantic Web to support automated high-end web applications.

This tutorial will introduce the semantic parsing task and will bring the audience up-to-date with the current research and state-of-the-art in semantic parsing. It will also provide insights about semantic parsing and how it relates to and differs from other natural language processing tasks. It will point out research challenges and some promising future directions for semantic parsing. The target audience will be NLP researchers and practitioners but no prior knowledge of semantic parsing will be assumed.

Rohit J. Kate is a postdoctoral fellow in the department of Computer Science at the University of Texas at Austin. He obtained his Ph.D. from the same place. His research interests are in natural language processing, especially in semantic parsing and information extraction, and in machine learning. He has worked extensively in semantic parsing, various forms of supervisions for semantic parser learners and kernel-based methods for natural language processing.

Yuk Wah Wong is a Senior Software Engineer at Google Pittsburgh. He obtained his Ph.D. from the University of Texas at Austin. His research interests are in natural language processing and machine learning. His thesis topic was on semantic parsing and generation using statistical machine translation techniques. Since joining Google, he has worked on information extraction, data integration, and natural language processing, with applications in web search and vertical search.

T5: Tree-based and Forest-based Translation

July 11, 09:00–12:30, Venue A, Hall X

Yang Liu and Liang Huang

The past several years have witnessed rapid advances in syntax-based machine translation, which exploits natural language syntax to guide translation. Depending on the type of input, most of these efforts can be divided into two broad categories: (a) string-based systems whose input is a string, which is simultaneously parsed and translated by a synchronous grammar (Wu, 1997; Chiang, 2005; Galley et al., 2006), and (b) tree-based systems whose input is already a parse tree to be directly converted into a target tree or string (Lin, 2004; Ding and Palmer, 2005; Quirk et al., 2005; Liu et al., 2006; Huang et al., 2006).

Compared with their string-based counterparts, tree-based systems offer many attractive features: they are much faster in decoding (linear time vs. cubic time), do not require sophisticated binarization (Zhang et al., 2006), and can use separate grammars for parsing and translation (e.g. a context-free grammar for the former and a tree substitution grammar for the latter).

However, despite these advantages, most tree-based systems suffer from a major drawback: they only use 1-best parse trees to direct translation, which potentially introduces translation mistakes due to parsing errors (Quirk and Corston-Oliver, 2006). This situation becomes worse for resource-poor source languages without enough Treebank data to train a high-accuracy parser.

This problem can be alleviated elegantly by using packed forests (Huang, 2008), which encodes exponentially many parse trees in a polynomial space. Forest-based systems (Mi et al., 2008; Mi and Huang, 2008) thus take a packed forest instead of

a parse tree as an input. In addition, packed forests could also be used for translation rule extraction, which helps alleviate the propagation of parsing errors into rule set. Forest-based translation can be regarded as a compromise between the string-based and tree-based methods, while combining the advantages of both: decoding is still fast, yet does not commit to a single parse. Surprisingly, translating a forest of millions of trees is even faster than translating 30 individual trees, and offers significantly better translation quality. This approach has since become a popular topic.

This tutorial surveys tree-based and forest-based translation methods. For each approach, we will discuss the two fundamental tasks: decoding, which performs the actual translation, and rule extraction, which learns translation rules from real-world data automatically. Finally, we will introduce some more recent developments to tree-based and forest-based translation, such as tree sequence based models, tree-to-tree models, joint parsing and translation, and faster decoding algorithms. We will conclude our talk by pointing out some directions for future work.

Yang Liu is an Associate Researcher at Institute of Computing Technology, Chinese Academy of Sciences (CAS/ICT). He obtained his PhD from CAS/ICT in 2007. His research interests include syntax-based translation, word alignment, and system combination. He has published five ACL full papers in the machine translation area in the recent five years. His work on "tree-to-string translation" received a Meritorious Asian NLP Paper Award at ACL 2006.

Liang Huang is a Computer Scientist at Information Sciences Institute, University of Southern California (USC/ISI), and a Research Assistant Professor at USC's Computer Science Dept. He obtained his PhD from the University of Pennsylvania under Aravind Joshi and Kevin Knight. His research interests are mainly in the theoretical aspects of NLP, esp. efficient algorithms in parsing and translation. His work on "forest-based algorithms" received an Outstanding Paper Award at ACL 2008, as well as Best Paper Nominations at ACL 2007 and EMNLP 2008. He has taught two tutorials on Advanced Dynamic Programming at COLING 2008 and NAACL 2009 and is currently (co-)teaching two NLP courses at USC.

T6: Discourse Structure: Theory, Practice and Use

July 11, 09:00–12:30, Venue A, Hall IV

Bonnie Webber, Markus Egg and Valia Kordoni

Discourse structure concerns the ways that discourses (monologic, dialogic and multi-party) are organised and those aspects of meaning that such organisation encodes. It is a potent influence on clause-level syntax, and the meaning it encodes is as essential to communication as that conveyed in a clause. Hence no modern language technology (LT) – information extraction, machine translation, opinion mining, or summarisation – can fully succeed without taking discourse structure into account. Attendees to this tutorial should gain insight into discourse structure (discourse relations; scope of attribution, modality and negation; centering; topic structure; dialogue moves and acts; macro-structure), its relevance for LT, and methods and resources that support its use. Our target audience are researchers and practitioners in LT (not necessarily discourse) who are interested in LT tasks that involve or could benefit from considering language and communication beyond the individual sentence.

Bonnie Webber is a Professor of Informatics at Edinburgh University. She is best known for work on Question Answering (starting with LUNAR in the early 70's) and discourse phenomena (starting with her PhD thesis on discourse anaphora). She has also carried out research on animation from instructions, medical decision support systems and biomedical text processing.

Markus Egg is a Professor of Linguistics at the Dept. of English and American Studies of the Humboldt University in Berlin. His main areas of interest are syntax, semantics, pragmatics, and discourse; the interfaces between them; and their implementation in NLP systems.

Valia Kordoni is a Senior Researcher at the Language Technology Lab of the German Research Centre for Artificial Intelligence (DFKI GmbH) and an assistant professor at the Department of Computational Linguistics of Saarland University. Her main areas of interest are syntax, semantics, pragmatics and discourse. She works on the theoretical development of these areas as well as on their implementation in NLP systems.

7 Main Conference, Day 1, July 12

Overview: Main Conference - Day 1 · Monday, July 12

7:00–8:45	Registration. Venue A, Foyer				
8:45–9:00 9:00–10:00	Opening. Venue A, Aula Invited talk				
10:00–10:30	*Coffee/Tea Break*				
	Venue A, Aula	Venue A, X	Venue A, IX	Venue B, 3	Venue B, 4
10:30–11:45	Parsing 1	Semantics 1	Spoken Language	Resources and Evaluation	Information Extraction 1
11:45–11:55	*Short Break*				
11:55–13:15	Short talks: Translation 1	Short talks: Discourse and Generation	Short talks: Psycho-linguistics, Resources and MT Evaluation	Short talks: Semantics 2	Short talks: Information Retrieval, Extraction, and Ontologies
13:15–15:00	Posters, Venue A, Foyer. *Lunch.*				
15:00–16:15	Translation 2	Parsing 2	Morphology	Sentiment 1	Selectional Preferences
16:15–16:45	*Coffee/Tea Break*				
16.45–18:00	Translation 3	Tagging	Grammar Formalisms	Summarization 1	Sentiment 2

Monday July 12, 2010

7:00–8:45 ***Registration***

Opening and Invited Talk

Venue A, Aula. Chair: Sandra Carberry (during the Invited Talk)

8:45–9:00 ***Opening***

9:00–10:00 ***Towards a Psycholinguistics of Social Interaction***

Zenzi M Griffin, University of Texas at Austin

Abstract: In studying spoken language production and comprehension, psycholinguistic researchers have typically designed experiments in which content consists of decontextualized utterances, narratives, or descriptions of visual displays (even when studying naïve participants in dialog). Like the drunk in the night who looks for keys where the light is brightest rather than where they were lost, we have studied language processing under the easiest circumstances to manipulate and control rather than study the speech acts and discourse functions that language use more often involves. I will argue that we now have resources available to extend experimental research to language use that has little or nothing to do with description. That is, psycholinguistics is ready to address language processing in interpersonal interactions. I will describe the results of a questionnaire study of parental name substitutions that led to this line of thought.

10:00–10:30 ***Coffee/Tea Break, Venue A, Foyer***

Parsing 1

Venue A, Aula. Chair: Jennifer Foster

10:30–10:55 *Efficient Third-Order Dependency Parsers*
Terry Koo and Michael Collins (p. 99)

10:55–11:20 *Dependency Parsing and Projection Based on Word-Pair Classification*
Wenbin Jiang and Qun Liu (p. 99)

11:20–11:45 *Bitext Dependency Parsing with Bilingual Subtree Constraints*
Wenliang Chen, Jun'ichi Kazama and Kentaro Torisawa (p. 99)

Semantics 1

Venue A, Hall X. Chair: Alexander Yates

10:30–10:55 *Computing Weakest Readings*
Alexander Koller and Stefan Thater (p. 99)

10:55–11:20 *Identifying Generic Noun Phrases*
Nils Reiter and Anette Frank (p. 100)

11:20–11:45 *Structural Semantic Relatedness: A Knowledge-Based Method to Named Entity Disambiguation*
Xianpei Han and Jun Zhao (p. 100)

Spoken Language

Venue A, Hall IX. Chair: Mikko Kurimo

10:30–10:55 *Correcting Errors in Speech Recognition with Articulatory Dynamics*
Frank Rudzicz (p. 100)

10:55–11:20 *Learning to Adapt to Unknown Users: Referring Expression Generation in Spoken Dialogue Systems*
Srinivasan Janarthanam and Oliver Lemon (p. 100)

11:20–11:45 *A Risk Minimization Framework for Extractive Speech Summarization*
Shih-Hsiang Lin and Berlin Chen (p. 101)

Resources and Evaluation

Venue B, Lecture Hall 3. Chair: Eduard Hovy

10:30–10:55 **Challenge Paper:** *The Human Language Project: Building a Universal Corpus of the World's Languages*
Steven Abney and Steven Bird (p. 101)

10:55–11:20 *Bilingual Lexicon Generation Using Non-Aligned Signatures*
Daphna Shezaf and Ari Rappoport (p. 101)

11:20–11:45 *Automatic Evaluation Method for Machine Translation Using Noun-Phrase Chunking*
Hiroshi Echizen-ya and Kenji Araki (p. 101)

Information Extraction 1

Venue B, Lecture Hall 4. Chair: Chin-Yew Lin

10:30–10:55 *Open Information Extraction Using Wikipedia*
Fei Wu and Daniel S. Weld (p. 102)

10:55–11:20 *SystemT: An Algebraic Approach to Declarative Information Extraction*
Laura Chiticariu, Rajasekar Krishnamurthy, Yunyao Li, Sriram Raghavan, Frederick Reiss and Shivakumar Vaithyanathan (p. 102)

11:20–11:45 **IBM Best Student Paper:** *Extracting Social Networks from Literary Fiction*
David Elson, Nicholas Dames and Kathleen McKeown (p. 102)

11:45–11:55 ***Short Break***

Short Talks: Translation 1

Venue A, Aula. Chair: Jörg Tiedemann

11:55–12:05 *Pseudo-Word for Phrase-Based Machine Translation*
Xiangyu Duan, Min Zhang and Haizhou Li (p. 102)

12:05–12:15 *Hierarchical Search for Word Alignment*
Jason Riesa and Daniel Marcu (p. 103)

12:15–12:25 *Paraphrase Lattice for Statistical Machine Translation*
Takashi Onishi, Masao Utiyama and Eiichiro Sumita (p. 103)

12:25–12:35 *A Joint Rule Selection Model for Hierarchical Phrase-Based Translation*
Lei Cui, Dongdong Zhang, Mu Li, Ming Zhou and Tiejun Zhao (p. 103)

12:35–12:45 *Learning Lexicalized Reordering Models from Reordering Graphs*
Jinsong Su, Yang Liu, Yajuan Lv, Haitao Mi and Qun Liu (p. 103)

12:45–12:55 *Filtering Syntactic Constraints for Statistical Machine Translation*
Hailong Cao and Eiichiro Sumita (p. 103)

12:55–13:05 *Diversify and Combine: Improving Word Alignment for Machine Translation on Low-Resource Languages*
Bing Xiang, Yonggang Deng and Bowen Zhou (p. 104)

13:05–13:15 *Efficient Path Counting Transducers for Minimum Bayes-Risk Decoding of Statistical Machine Translation Lattices*
Graeme Blackwood, Adrià de Gispert and William Byrne (p. 104)

Short Talks: Discourse and Generation

Venue A, Hall X. Chair: Oliver Lemon

11:55–12:05 *"Was It Good? It Was Provocative." Learning the Meaning of Scalar Adjectives*
Marie-Catherine de Marneffe, Christopher D. Manning and Christopher Potts (p. 104)

12:05–12:15 *The Same-Head Heuristic for Coreference*
Micha Elsner and Eugene Charniak (p. 104)

12:15–12:25 *Authorship Attribution Using Probabilistic Context-Free Grammars*
Sindhu Raghavan, Adriana Kovashka and Raymond Mooney (p. 104)

12:25–12:35 *The Impact of Interpretation Problems on Tutorial Dialogue*
Myroslava O. Dzikovska, Johanna D. Moore, Natalie Steinhauser and Gwendolyn Campbell (p. 105)

12:35–12:45 *Importance-Driven Turn-Bidding for Spoken Dialogue Systems*
Ethan Selfridge and Peter Heeman (p. 105)

12:45–12:55 *The Prevalence of Descriptive Referring Expressions in News and Narrative*
Raquel Hervas and Mark Finlayson (p. 105)

12:55–13:05 *Preferences versus Adaptation during Referring Expression Generation*
Martijn Goudbeek and Emiel Krahmer (p. 105)

13:05–13:15 *Entity-Based Local Coherence Modelling Using Topological Fields*
Jackie Chi Kit Cheung and Gerald Penn (p. 105)

Short Talks: Psycholinguistics, Resources and MT Evaluation

Venue A, Hall IX. Chair: Amit Dubey

11:55–12:05 **Challenge Paper:** *Cognitively Plausible Models of Human Language Processing*
Frank Keller (p. 106)

12:05–12:15 *Syntactic and Semantic Factors in Processing Difficulty: An Integrated Measure*
Jeff Mitchell, Mirella Lapata, Vera Demberg and Frank Keller (p. 106)

12:15–12:25 *The Manually Annotated Sub-Corpus: A Community Resource for and by the People*
Nancy Ide, Collin Baker, Christiane Fellbaum and Rebecca Passonneau (p. 106)

12:25–12:35 *Correcting Errors in a Treebank Based on Synchronous Tree Substitution Grammar*
Yoshihide Kato and Shigeki Matsubara (p. 106)

12:35–12:45 *Rebanking CCGbank for Improved NP Interpretation*
Matthew Honnibal, James R. Curran and Johan Bos (p. 106)

12:45–12:55 *BabelNet: Building a Very Large Multilingual Semantic Network*
Roberto Navigli and Simone Paolo Ponzetto (p. 107)

12:55–13:05 *Evaluating Machine Translations Using mNCD*
Marcus Dobrinkat, Tero Tapiovaara, Jaakko Väyrynen and Kimmo Kettunen (p. 107)

13:05–13:15 *Tackling Sparse Data Issue in Machine Translation Evaluation*
Ondřej Bojar, Kamil Kos and David Mareček (p. 107)

Short Talks: Semantics 2

Venue B, Lecture Hall 3. Chair: Manfred Pinkal

11:55–12:05 *Exemplar-Based Models for Word Meaning in Context*
Katrin Erk and Sebastian Pado (p. 107)

12:05–12:15 *Fully Unsupervised Core-Adjunct Argument Classification*
Omri Abend and Ari Rappoport (p. 107)

12:15–12:25 *A Structured Model for Joint Learning of Argument Roles and Predicate Senses*
Yotaro Watanabe, Masayuki Asahara and Yuji Matsumoto (p. 107)

12:25–12:35 *Semantics-Driven Shallow Parsing for Chinese Semantic Role Labeling*
Weiwei Sun (p. 108)

12:35–12:45 *Towards Open-Domain Semantic Role Labeling*
Danilo Croce, Cristina Giannone, Paolo Annesi and Roberto Basili (p. 108)

12:45–12:55 *Collocation Extraction beyond the Independence Assumption*
Gerlof Bouma (p. 108)

12:55–13:05 *Automatic Collocation Suggestion in Academic Writing*
Jian-Cheng Wu, Yu-Chia Chang, Teruko Mitamura and Jason S. Chang (p. 108)

13:05–13:15 *A Bayesian Method for Robust Estimation of Distributional Similarities*
Jun'ichi Kazama, Stijn De Saeger, Kow Kuroda, Masaki Murata and Kentaro Torisawa (p. 108)

Short Talks: Information Retrieval, Extraction, and Ontologies

Venue B, Lecture Hall 4. Chair: Pushpak Bhattacharyya

11:55–12:05 *Recommendation in Internet Forums and Blogs*
Jia Wang, Qing Li, Yuanzhu Peter Chen and Zhangxi Lin (p. 109)

12:05–12:15 *Event-Based Hyperspace Analogue to Language for Query Expansion*
Tingxu Yan, Tamsin Maxwell, Dawei Song, Yuexian Hou and Peng Zhang (p. 109)

12:15–12:25 *Learning Phrase-Based Spelling Error Models from Clickthrough Data*
Xu Sun, Jianfeng Gao, Daniel Micol and Chris Quirk (p. 109)

12:25–12:35 *Inducing Domain-Specific Semantic Class Taggers from (Almost) Nothing*
Ruihong Huang and Ellen Riloff (p. 109)

12:35–12:45 *Learning 5000 Relational Extractors*
Raphael Hoffmann, Congle Zhang and Daniel S. Weld (p. 110)

12:45–12:55 *Unsupervised Ontology Induction from Text*
Hoifung Poon and Pedro Domingos (p. 110)

12:55–13:05 *Automatically Generating Term Frequency Induced Taxonomies*
Karin Murthy, Tanveer A Faruquie, L Venkata Subramaniam, Hima Prasad K and Mukesh Mohania (p. 110)

13:05–13:15 *Complexity Assumptions in Ontology Verbalisation*
Richard Power (p. 110)

Poster Session

Venue A, Foyer

13:15–15:00 ***Poster Presentations and Lunch (Complimentary)***

Posters: Translation

(01) *Pseudo-Word for Phrase-Based Machine Translation*
Xiangyu Duan, Min Zhang and Haizhou Li (p. 102)

(02) *Hierarchical Search for Word Alignment*
Jason Riesa and Daniel Marcu (p. 103)

(03) *Paraphrase Lattice for Statistical Machine Translation*
Takashi Onishi, Masao Utiyama and Eiichiro Sumita (p. 103)

(04) *A Joint Rule Selection Model for Hierarchical Phrase-Based Translation*
Lei Cui, Dongdong Zhang, Mu Li, Ming Zhou and Tiejun Zhao (p. 103)

(05) *Learning Lexicalized Reordering Models from Reordering Graphs*
Jinsong Su, Yang Liu, Yajuan Lv, Haitao Mi and Qun Liu (p. 103)

(06) *Filtering Syntactic Constraints for Statistical Machine Translation*
Hailong Cao and Eiichiro Sumita (p. 103)

(07) *Diversify and Combine: Improving Word Alignment for Machine Translation on Low-Resource Languages*
Bing Xiang, Yonggang Deng and Bowen Zhou (p. 104)

(08) *Efficient Path Counting Transducers for Minimum Bayes-Risk Decoding of Statistical Machine Translation Lattices*
Graeme Blackwood, Adrià de Gispert and William Byrne (p. 104)

(09) *Word Alignment with Synonym Regularization*
Hiroyuki Shindo, Akinori Fujino and Masaaki Nagata (p. 110)

(10) *Better Filtration and Augmentation for Hierarchical Phrase-Based Translation Rules*
Zhiyang Wang, Yajuan Lv, Qun Liu and Young-Sook Hwang (p. 111)

(11) *Fixed Length Word Suffix for Factored Statistical Machine Translation*
Narges Sharif Razavian and Stephan Vogel (p. 111)

Posters: Generation

(13) *The Prevalence of Descriptive Referring Expressions in News and Narrative*
Raquel Hervas and Mark Finlayson (p. 105)

(14) *Preferences versus Adaptation during Referring Expression Generation*
Martijn Goudbeek and Emiel Krahmer (p. 105)

(15) *Entity-Based Local Coherence Modelling Using Topological Fields*
Jackie Chi Kit Cheung and Gerald Penn (p. 105)

Posters: Information Retrieval and Extraction

(16) *Recommendation in Internet Forums and Blogs*
Jia Wang, Qing Li, Yuanzhu Peter Chen and Zhangxi Lin (p. 109)

(17) *Event-Based Hyperspace Analogue to Language for Query Expansion*
Tingxu Yan, Tamsin Maxwell, Dawei Song, Yuexian Hou and Peng Zhang (p. 109)

(18) *Learning Phrase-Based Spelling Error Models from Clickthrough Data*
Xu Sun, Jianfeng Gao, Daniel Micol and Chris Quirk (p. 109)

(19) *Inducing Domain-Specific Semantic Class Taggers from (Almost) Nothing*
Ruihong Huang and Ellen Riloff (p. 109)

(20) *Learning 5000 Relational Extractors*
Raphael Hoffmann, Congle Zhang and Daniel S. Weld (p. 110)

Posters: Discourse

(21) *"Was It Good? It Was Provocative." Learning the Meaning of Scalar Adjectives*
Marie-Catherine de Marneffe, Christopher D. Manning and Christopher Potts (p. 104)

(22) *The Same-Head Heuristic for Coreference*
Micha Elsner and Eugene Charniak (p. 104)

(23) *Authorship Attribution Using Probabilistic Context-Free Grammars*
Sindhu Raghavan, Adriana Kovashka and Raymond Mooney (p. 104)

(24) *The Impact of Interpretation Problems on Tutorial Dialogue*
Myroslava O. Dzikovska, Johanna D. Moore, Natalie Steinhauser and Gwendolyn Campbell (p. 105)

(25) *Importance-Driven Turn-Bidding for Spoken Dialogue Systems*
Ethan Selfridge and Peter Heeman (p. 105)

(44) *A Bayesian Method for Robust Estimation of Distributional Similarities*
Jun'ichi Kazama, Stijn De Saeger, Kow Kuroda, Masaki Murata and Kentaro Torisawa (p. 108)

(45) *Improving Chinese Semantic Role Labeling with Rich Syntactic Features*
Weiwei Sun (p. 112)

Posters: Ontologies

(46) *Unsupervised Ontology Induction from Text*
Hoifung Poon and Pedro Domingos (p. 110)

(47) *Automatically Generating Term Frequency Induced Taxonomies*
Karin Murthy, Tanveer A Faruquie, L Venkata Subramaniam, Hima Prasad K and Mukesh Mohania (p. 110)

(48) *Complexity Assumptions in Ontology Verbalisation*
Richard Power (p. 110)

Posters: Psycholinguistics

(49) *Cognitively Plausible Models of Human Language Processing*
Frank Keller (p. 106)

(50) *Syntactic and Semantic Factors in Processing Difficulty: An Integrated Measure*
Jeff Mitchell, Mirella Lapata, Vera Demberg and Frank Keller (p. 106)

Translation 2

Venue A, Aula. Chair: Kemal Oflazer

15:00–15:25 *Exploring Syntactic Structural Features for Sub-Tree Alignment Using Bilingual Tree Kernels*
Jun Sun, Min Zhang and Chew Lim Tan (p. 112)

15:25–15:50 *Discriminative Pruning for Discriminative ITG Alignment*
Shujie Liu, Chi-Ho Li and Ming Zhou (p. 112)

15:50–16:15 *Fine-Grained Tree-to-String Translation Rule Extraction*
Xianchao Wu, Takuya Matsuzaki and Jun'ichi Tsujii (p. 112)

Parsing 2

Venue A, Hall X. Chair: Josef van Genabith

15:00–15:25 *Accurate Context-Free Parsing with Combinatory Categorial Grammar*
Timothy A. D. Fowler and Gerald Penn (p. 113)

15:25–15:50 *Faster Parsing by Supertagger Adaptation*
Jonathan K. Kummerfeld, Jessika Roesner, Tim Dawborn, James Haggerty, James R. Curran and Stephen Clark (p. 113)

15:50–16:15 *Using Smaller Constituents Rather Than Sentences in Active Learning for Japanese Dependency Parsing*
Manabu Sassano and Sadao Kurohashi (p. 113)

Morphology

Venue A, Hall IX. Chair: Markus Dickinson

15:00–15:25 *Conditional Random Fields for Word Hyphenation*
Nikolaos Trogkanis and Charles Elkan (p. 113)

15:25–15:50 *Enhanced Word Decomposition by Calibrating the Decision Threshold of Probabilistic Models and Using a Model Ensemble*
Sebastian Spiegler and Peter A. Flach (p. 114)

15:50–16:15 *Word Representations: A Simple and General Method for Semi-Supervised Learning*
Joseph Turian, Lev-Arie Ratinov and Yoshua Bengio (p. 114)

Sentiment 1

Venue B, Lecture Hall 3. Chair: Christopher Pal

15:00–15:25 *Identifying Text Polarity Using Random Walks*
Ahmed Hassan and Dragomir Radev (p. 114)

15:25–15:50 *Sentiment Learning on Product Reviews via Sentiment Ontology Tree*
Wei Wei and Jon Atle Gulla (p. 114)

15:50–16:15 *Employing Personal/Impersonal Views in Supervised and Semi-Supervised Sentiment Classification*
Shoushan Li, Chu-Ren Huang, Guodong Zhou and Sophia Yat Mei Lee (p. 115)

Selectional Preferences

Venue B, Lecture Hall 4. Chair: Anette Frank

15:00–15:25 *A Latent Dirichlet Allocation Method for Selectional Preferences*
Alan Ritter, Mausam and Oren Etzioni (p. 115)

15:25–15:50 *Latent Variable Models of Selectional Preference*
Diarmuid Ó Séaghdha (p. 115)

15:50–16:15 *Improving the Use of Pseudo-Words for Evaluating Selectional Preferences*
Nathanael Chambers and Daniel Jurafsky (p. 115)

16:15–16:45 ***Coffee/Tea Break, Venue A and B, Foyer***

Translation 3

Venue A, Aula. Chair: Adrià de Gispert

16:45–17:10 *Syntax-to-Morphology Mapping in Factored Phrase-Based Statistical Machine Translation from English to Turkish*
Reyyan Yeniterzi and Kemal Oflazer (p. 116)

17:10–17:35 *Hindi-to-Urdu Machine Translation through Transliteration*
Nadir Durrani, Hassan Sajjad, Alexander Fraser and Helmut Schmid (p. 116)

17:35–18:00 *Training Phrase Translation Models with Leaving-One-Out*
Joern Wuebker, Arne Mauser and Hermann Ney (p. 116)

Tagging

Venue A, Hall X. Chair: Hoifung Poon

16:45–17:10 *Efficient Staggered Decoding for Sequence Labeling*
Nobuhiro Kaji, Yasuhiro Fujiwara, Naoki Yoshinaga and Masaru Kitsuregawa (p. 117)

17:10–17:35 *Minimized Models and Grammar-Informed Initialization for Supertagging with Highly Ambiguous Lexicons*
Sujith Ravi, Jason Baldridge and Kevin Knight (p. 117)

17:35–18:00 *Practical Very Large Scale CRFs*
Thomas Lavergne, Olivier Cappé and François Yvon (p. 117)

Grammar Formalisms

Venue A, Hall IX. Chair: Gerald Penn

16:45–17:10 **Survey Paper:** *On the Computational Complexity of Dominance Links in Grammatical Formalisms*
Sylvain Schmitz (p. 117)

17:10–17:35 *Optimal Rank Reduction for Linear Context-Free Rewriting Systems with Fan-Out Two*
Benoît Sagot and Giorgio Satta (p. 117)

17:35–18:00 *The Importance of Rule Restrictions in CCG*
Marco Kuhlmann, Alexander Koller and Giorgio Satta (p. 118)

Summarization 1

Venue B, Lecture Hall 3. Chair: Xiaojun Wan

16:45–17:10 *Automatic Evaluation of Linguistic Quality in Multi-Document Summarization*
Emily Pitler, Annie Louis and Ani Nenkova (p. 118)

17:10–17:35 *Identifying Non-Explicit Citing Sentences for Citation-Based Summarization.*
Vahed Qazvinian and Dragomir Radev (p. 118)

17:35–18:00 *Automatic Generation of Story Highlights*
Kristian Woodsend and Mirella Lapata (p. 118)

Sentiment 2

Venue B, Lecture Hall 4. Chair: Georgios Paltoglou

16:45–17:10 *Sentence and Expression Level Annotation of Opinions in User-Generated Discourse*
Cigdem Toprak, Niklas Jakob and Iryna Gurevych (p. 119)

17:10–17:35 *Generating Focused Topic-Specific Sentiment Lexicons*
Valentin Jijkoun, Maarten de Rijke and Wouter Weerkamp (p. 119)

17:35–18:00 *Evaluating Multilanguage-Comparability of Subjectivity Analysis Systems*
Jungi Kim, Jin-Ji Li and Jong-Hyeok Lee (p. 119)

Overview: Main Conference, Day 2 · Tuesday, July 13

7:30–9:00	Registration. Venue A, Foyer					
9:00–10.00	Lifetime Achievement Award. Venue A, Aula					
10:00–10:30	*Coffee/Tea Break*					
	Venue A, Aula	Venue A, X	Venue A, IX	Venue B, 3	Venue B, 4	
10:30–11:45	Translation 4	Information Extraction 2	Student Research Workshop	Resources	Discourse 1	
11:45–11:55	*Short Break*					
11:55–13:15	Short talks: Translation and Parsing	Short talks: Machine Learning and Statistical Methods	Short talks: Question Answering, Entailment and Sentiment	Short talks: Morphology and Information Extraction	Short talks: Speech, Multimodal, and Summa-rization	Venue A, VIII Student Research Workshop Posters
13:15–15:00	Posters, Venue A, Foyer. *Lunch.*					
15:00–16:15	Translation and Multilinguality	Machine Learning	Language Learning and Models of Language	Summar-ization 2	Semantics 3	Venue A, XI Software Demonst-rations
16:15–16:45	*Coffee/Tea Break*					
16.45–17:35	Semantics 4	Dialogue	Historical Linguistics	Decipherment	Tree Transducers	

Tuesday July 13, 2010

7:30–9:00 ***Registration***

Lifetime Achievement Award

Venue A, Aula. Chair: Ido Dagan

9:00–10:00 ***Lifetime Achievement Award Ceremony***

10:00–10:30 ***Coffee/Tea Break, Venue A, Foyer***

Translation 4

Venue A, Aula. Chair: Haifeng Wang

10:30–10:55 *Error Detection for Statistical Machine Translation Using Linguistic Features*
Deyi Xiong, Min Zhang and Haizhou Li (p. 120)

10:55–11:20 *TrustRank: Inducing Trust in Automatic Translations via Ranking*
Radu Soricut and Abdessamad Echihabi (p. 120)

11:20–11:45 *Bridging SMT and TM with Translation Recommendation*
Yifan He, Yanjun Ma, Josef van Genabith and Andy Way (p. 120)

Information Extraction 2

Venue A, Hall X. Chair: Nianwen Xue

10:30–10:55 *On Jointly Recognizing and Aligning Bilingual Named Entities*
Yufeng Chen, Chengqing Zong and Keh-Yih Su (p. 120)

10:55–11:20 *Generating Templates of Entity Summaries with an Entity-Aspect Model and Pattern Mining*
Peng Li, Jing Jiang and Yinglin Wang (p. 121)

11:20–11:45 *Comparable Entity Mining from Comparative Questions*
Shasha Li, Chin-Yew Lin, Young-In Song and Zhoujun Li (p. 121)

Student Research Workshop

Venue A, Hall IX. Chair: Jan Raab

10:30–10:40 *Non-Cooperation in Dialogue*
Brian Plüss (p. 121)

10:40–10:50 *Towards Relational POMDPs for Adaptive Dialogue Management*
Pierre Lison (p. 121)

10:50–11:00 *WSD as a Distributed Constraint Optimization Problem*
Siva Reddy and Abhilash Inumella (p. 122)

11:00–11:10 *A Probabilistic Generative Model for an Intermediate Constituency-Dependency Representation*
Federico Sangati (p. 122)

11:10–11:20 *Sentiment Translation through Lexicon Induction*
Christian Scheible (p. 122)

11:20–11:30 *Unsupervised Search for The Optimal Segmentation for Statistical Machine Translation*
Coşkun Mermer and Ahmet Afşın Akın (p. 122)

11:30–11:40 *How Spoken Language Corpora can Refine Current Speech Motor Training Methodologies*
Daniil Umanski and Federico Sangati (p. 123)

Resources

Venue B, Lecture Hall 3. Chair: Nancy Ide

10:30–10:55 *Towards Robust Multi-Tool Tagging. An OWL/DL-Based Approach*
Christian Chiarcos (p. 123)

10:55–11:20 *Temporal Information Processing of a New Language: Fast Porting with Minimal Resources*
Francisco Costa and António Branco (p. 123)

11:20–11:45 *A Taxonomy, Dataset, and Classifier for Automatic Noun Compound Interpretation*
Stephen Tratz and Eduard Hovy (p. 123)

Discourse 1

Venue B, Lecture Hall 4. Chair: Peter Heeman

10:30–10:55 **Survey Paper:** *Models of Metaphor in NLP*
Ekaterina Shutova (p. 123)

10:55–11:20 *A Game-Theoretic Model of Metaphorical Bargaining*
Beata Beigman Klebanov and Eyal Beigman (p. 124)

11:20–11:45 *Kernel Based Discourse Relation Recognition with Temporal Ordering Information*
WenTing Wang, Jian Su and Chew Lim Tan (p. 124)

11:45–11:55 ***Short Break***

Short Talks: Translation and Parsing

Venue A, Aula. Chair: Julia Hockenmaier

11:55–12:05 *Balancing User Effort and Translation Error in Interactive Machine Translation via Confidence Measures*
Jesús González Rubio, Daniel Ortiz Martínez and Francisco Casacuberta (p. 124)

12:05–12:15 *Improving Arabic-to-English Statistical Machine Translation by Reordering Post-Verbal Subjects for Alignment*
Marine Carpuat, Yuval Marton and Nizar Habash (p. 124)

12:15–12:25 *Learning Common Grammar from Multilingual Corpus*
Tomoharu Iwata, Daichi Mochihashi and Hiroshi Sawada (p. 124)

12:25–12:35 *Hierarchical Joint Learning: Improving Joint Parsing and Named Entity Recognition with Non-Jointly Labeled Data*
Jenny Rose Finkel and Christopher D. Manning (p. 125)

12:35–12:45 *Detecting Errors in Automatically-Parsed Dependency Relations*
Markus Dickinson (p. 125)

12:45–12:55 *Tree-Based Deterministic Dependency Parsing — An Application to Nivre's Method —*
Kotaro Kitagawa and Kumiko Tanaka-Ishii (p. 125)

12:55–13:05 *Sparsity in Dependency Grammar Induction*
Jennifer Gillenwater, Kuzman Ganchev, João Graça, Fernando Pereira and Ben Taskar (p. 125)

13:05–13:15 *Top-Down K-Best A* Parsing*
Adam Pauls, Dan Klein and Chris Quirk (p. 126)

Short Talks: Machine Learning and Statistical Methods

Venue A, Hall X. Chair: Dekang Lin

11:55–12:05 *Simple Semi-Supervised Training of Part-Of-Speech Taggers*
Anders Søgaard (p. 126)

12:05–12:15 *Efficient Optimization of an MDL-Inspired Objective Function for Unsupervised Part-Of-Speech Tagging*
Ashish Vaswani, Adam Pauls and David Chiang (p. 126)

12:15–12:25 **Best Short Paper:** *SVD and Clustering for Unsupervised POS Tagging*
Michael Lamar, Yariv Maron, Mark Johnson and Elie Bienenstock (p. 126)

12:25–12:35 *Intelligent Selection of Language Model Training Data*
Robert C. Moore and William Lewis (p. 126)

12:35–12:45 *Blocked Inference in Bayesian Tree Substitution Grammars*
Trevor Cohn and Phil Blunsom (p. 127)

12:45–12:55 *Boosting-Based System Combination for Machine Translation*
Tong Xiao, Jingbo Zhu, Muhua Zhu and Huizhen Wang (p. 127)

12:55–13:05 *Fine-Grained Genre Classification Using Structural Learning Algorithms*
Zhili Wu, Katja Markert and Serge Sharoff (p. 127)

13:05–13:15 *Online Generation of Locality Sensitive Hash Signatures*
Benjamin Van Durme and Ashwin Lall (p. 127)

Short Talks: Question Answering, Entailment and Sentiment

Venue A, Hall IX. Chair: Sanda Harabagiu

11:55–12:05 *Metadata-Aware Measures for Answer Summarization in Community Question Answering*
Mattia Tomasoni and Minlie Huang (p. 128)

12:05–12:15 *Optimizing Question Answering Accuracy by Maximizing Log-Likelihood*
Matthias H. Heie, Edward W. D. Whittaker and Sadaoki Furui (p. 128)

12:15–12:25 *Generating Entailment Rules from FrameNet*
Roni Ben Aharon, Idan Szpektor and Ido Dagan (p. 128)

12:25–12:35 *Don't 'Have a Clue'? Unsupervised Co-Learning of Downward-Entailing Operators.*
Cristian Danescu-Niculescu-Mizil and Lillian Lee (p. 128)

12:35–12:45 *Vocabulary Choice as an Indicator of Perspective*
Beata Beigman Klebanov, Eyal Beigman and Daniel Diermeier (p. 128)

12:45–12:55 *Cross Lingual Adaptation: An Experiment on Sentiment Classifications*
Bin Wei and Christopher Pal (p. 128)

12:55–13:05 *Using Anaphora Resolution to Improve Opinion Target Identification in Movie Reviews*
Niklas Jakob and Iryna Gurevych (p. 129)

13:05–13:15 *Hierarchical Sequential Learning for Extracting Opinions and Their Attributes*
Yejin Choi and Claire Cardie (p. 129)

Short Talks: Morphology and Information Extraction

Venue B, Lecture Hall 3. Chair: Gosse Bouma

11:55–12:05 *A Hybrid Rule/Model-Based Finite-State Framework for Normalizing SMS Messages*
Richard Beaufort, Sophie Roekhaut, Louise-Amélie Cougnon and Cédrick Fairon (p. 129)

12:05–12:15 *Letter-Phoneme Alignment: An Exploration*
Sittichai Jiampojamarn and Grzegorz Kondrak (p. 130)

12:15–12:25 *Jointly Optimizing a Two-Step Conditional Random Field Model for Machine Transliteration and Its Fast Decoding Algorithm*
Dong Yang, Paul Dixon and Sadaoki Furui (p. 130)

12:25–12:35 *Arabic Named Entity Recognition: Using Features Extracted from Noisy Data*
Yassine Benajiba, Imed Zitouni, Mona Diab and Paolo Rosso (p. 130)

12:35–12:45 *Extracting Sequences from the Web*
Anthony Fader, Stephen Soderland and Oren Etzioni (p. 130)

12:45–12:55 *An Entity-Level Approach to Information Extraction*
Aria Haghighi and Dan Klein (p. 130)

12:55–13:05 *Using Document Level Cross-Event Inference to Improve Event Extraction*
Shasha Liao and Ralph Grishman (p. 130)

13:05–13:15 *A Semi-Supervised Key Phrase Extraction Approach: Learning from Title Phrases through a Document Semantic Network*
Decong Li, Sujian Li, Wenjie Li, Wei Wang and Weiguang Qu (p. 131)

Short Talks: Speech, Multimodal, and Summarization

Venue B, Lecture Hall 4. Chair: Berlin Chen

11:55–12:05 *Domain Adaptation of Maximum Entropy Language Models*
Tanel Alumäe and Mikko Kurimo (p. 131)

12:05–12:15 *Decision Detection Using Hierarchical Graphical Models*
Trung H. Bui and Stanley Peters (p. 131)

12:15–12:25 *Now, Where Was I? Resumption Strategies for an In-Vehicle Dialogue System*
Jessica Villing (p. 131)

12:25–12:35 *Using Speech to Reply to SMS Messages While Driving: An In-Car Simulator User Study*
Yun-Cheng Ju and Tim Paek (p. 132)

12:35–12:45 *Learning to Follow Navigational Directions*
Adam Vogel and Daniel Jurafsky (p. 132)

12:45–12:55 *Classification of Feedback Expressions in Multimodal Data*
Costanza Navarretta and Patrizia Paggio (p. 132)

12:55–13:05 *A Hybrid Hierarchical Model for Multi-Document Summarization*
Asli Celikyilmaz and Dilek Hakkani-Tur (p. 132)

13:05–13:15 *Optimizing Informativeness and Readability for Sentiment Summarization*
Hitoshi Nishikawa, Takaaki Hasegawa, Yoshihiro Matsuo and Genichiro Kikui (p. 132)

Student Research Workshop Poster Session

Venue A, Room VIII

11:55–13:15 ***Poster Presentations***

(53) *Mood Patterns and Affective Lexicon Access in Weblogs*
Thin Nguyen (p. 133)

(54) *Growing Related Words from Seed via User Behaviors: A Re-ranking Based Approach*
Yabin Zheng, Zhiyuan Liu and Lixing Xie (p. 133)

(55) *Transition-Based Parsing with Confidence-Weighted Classification*
Martin Haulrich (p. 133)

(56) *Expanding Verb Coverage in Cyc With VerbNet*
Clifton McFate (p. 133)

(57) *A Framework for Figurative Language Detection Based on Sense Differentiation*
Daria Bogdanova (p. 133)

(58) *Automatic Selectional Preference Acquisition for Latin verbs*
Barbara McGillivray (p. 134)

(59) *Edit Tree Distance Alignments for Semantic Role Labelling*
Hector-Hugo Franco-Penya (p. 134)

(60) *Automatic Sanskrit Segmentizer Using Finite State Transducers*
Vipul Mittal (p. 134)

(61) *Adapting Self-training for Semantic Role Labeling*
Rasoul Samad Zadeh Kaljahi (p. 134)

(62) *Weakly Supervised Learning of Presupposition Relations between Verbs*
Galina Tremper (p. 134)

(63) *Importance of Linguistic Constraints in Statistical Dependency Parsing*
Bharat Ram Ambati (p. 134)

(64) *The Use of Formal Language Models in the Typology of the Morphology of Amerindian Languages*
Andres Osvaldo Porta (p. 135)

(65) *Non-Cooperation in Dialogue*
Brian Plüss (p. 121)

(66) *Towards Relational POMDPs for Adaptive Dialogue Management*
Pierre Lison (p. 121)

(67) *WSD as a Distributed Constraint Optimization Problem*
Siva Reddy and Abhilash Inumella (p. 122)

(68) *A Probabilistic Generative Model for an Intermediate Constituency-Dependency Representation*
Federico Sangati (p. 122)

(69) *Sentiment Translation through Lexicon Induction*
Christian Scheible (p. 122)

(70) *Unsupervised Search for The Optimal Segmentation for Statistical Machine Translation*
Coşkun Mermer and Ahmet Afşın Akın (p. 122)

(71) *How Spoken Language Corpora can Refine Current Speech Motor Training Methodologies*
Daniil Umanski and Federico Sangati (p. 123)

Poster Session

Venue A, Foyer

13:15–15:00 ***Poster Presentations and Lunch (Complimentary)***

Posters: Question Answering and Entailment

(01) *Metadata-Aware Measures for Answer Summarization in Community Question Answering*
Mattia Tomasoni and Minlie Huang (p. 128)

(02) *Optimizing Question Answering Accuracy by Maximizing Log-Likelihood*
Matthias H. Heie, Edward W. D. Whittaker and Sadaoki Furui (p. 128)

(03) *Generating Entailment Rules from FrameNet*
Roni Ben Aharon, Idan Szpektor and Ido Dagan (p. 128)

(04) *Don't 'Have a Clue'? Unsupervised Co-Learning of Downward-Entailing Operators.*
Cristian Danescu-Niculescu-Mizil and Lillian Lee (p. 128)

Posters: Sentiment

(05) *Vocabulary Choice as an Indicator of Perspective*
Beata Beigman Klebanov, Eyal Beigman and Daniel Diermeier (p. 128)

Posters: Machine Learning and Statistical Methods

Posters: Summarization

Translation and Multilinguality

Venue A, Aula. Chair: Marine Carpuat

15:00–15:25 *Improving Statistical Machine Translation with Monolingual Collocation*
Zhanyi Liu, Haifeng Wang, Hua Wu and Sheng Li (p. 137)

15:25–15:50 *Bilingual Sense Similarity for Statistical Machine Translation*
Boxing Chen, George Foster and Roland Kuhn (p. 137)

15:50–16:15 *Untangling the Cross-Lingual Link Structure of Wikipedia*
Gerard de Melo and Gerhard Weikum (p. 137)

Machine Learning

Venue A, Hall X. Chair: Joseph Turian

15:00–15:25 *Bucking the Trend: Large-Scale Cost-Focused Active Learning for Statistical Machine Translation*
Michael Bloodgood and Chris Callison-Burch (p. 138)

15:25–15:50 *Creating Robust Supervised Classifiers via Web-Scale N-Gram Data*
Shane Bergsma, Emily Pitler and Dekang Lin (p. 138)

15:50–16:15 *Convolution Kernel over Packed Parse Forest*
Min Zhang, Hui Zhang and Haizhou Li (p. 138)

Language Learning and Models of Language

Venue A, Hall IX. Chair: Alexander Clark

15:00–15:25 *Estimating Strictly Piecewise Distributions*
Jeffrey Heinz and James Rogers (p. 138)

15:25–15:50 *String Extension Learning*
Jeffrey Heinz (p. 138)

15:50–16:15 *Compositional Matrix-Space Models of Language*
Sebastian Rudolph and Eugenie Giesbrecht (p. 138)

Summarization 2

Venue B, Lecture Hall 3. Chair: Bonnie Webber

15:00–15:25 *Cross-Language Document Summarization Based on Machine Translation Quality Prediction*
Xiaojun Wan, Huiying Li and Jianguo Xiao (p. 139)

15:25–15:50 *A New Approach to Improving Multilingual Summarization Using a Genetic Algorithm*
Marina Litvak, Mark Last and Menahem Friedman (p. 139)

15:50–16:15 *Bayesian Synchronous Tree-Substitution Grammar Induction and Its Application to Sentence Compression*
Elif Yamangil and Stuart M. Shieber (p. 139)

Semantics 3

Venue B, Lecture Hall 4. Chair: Katrin Erk

15:00–15:25 *Contextualizing Semantic Representations Using Syntactically Enriched Vector Models*
Stefan Thater, Hagen Fürstenau and Manfred Pinkal (p. 139)

15:25–15:50 *Bootstrapping Semantic Analyzers from Non-Contradictory Texts*
Ivan Titov and Mikhail Kozhevnikov (p. 140)

15:50–16:15 *Open-Domain Semantic Role Labeling by Modeling Word Spans*
Fei Huang and Alexander Yates (p. 140)

Software Demonstration Session

Venue A, Room XI

15:00–17:35 ***Software Demonstrations***

(73) *Grammar Prototyping and Testing with the LinGO Grammar Matrix Customization System*
Emily M. Bender, Scott Drellishak, Antske Fokkens, Michael Wayne Goodman, Daniel P. Mills, Laurie Poulson and Safiyyah Saleem (p. 140)

(74) *cdec: A Decoder, Alignment, and Learning Framework for Finite-State and Context-Free Translation Models*
Chris Dyer, Adam Lopez, Juri Ganitkevitch, Jonathan Weese, Ferhan Ture, Phil Blunsom, Hendra Setiawan, Vladimir Eidelman and Philip Resnik (p. 140)

(75) *Beetle II: A System for Tutoring and Computational Linguistics Experimentation*
Myroslava O. Dzikovska, Johanna D. Moore, Natalie Steinhauser, Gwendolyn Campbell, Elaine Farrow and Charles B. Callaway (p. 141)

(76) *GernEdiT - The GermaNet Editing Tool*
Verena Henrich and Erhard Hinrichs (p. 141)

(77) *WebLicht: Web-Based LRT Services for German*
Erhard Hinrichs, Marie Hinrichs and Thomas Zastrow (p. 141)

(78) *The S-Space Package: An Open Source Package for Word Space Models*
David Jurgens and Keith Stevens (p. 141)

(79) *Talking NPCs in a Virtual Game World*
Tina Klüwer, Peter Adolphs, Feiyu Xu, Hans Uszkoreit and Xiwen Cheng (p. 141)

(80) *An Open-Source Package for Recognizing Textual Entailment*
Milen Kouylekov and Matteo Negri (p. 142)

(81) *Personalising Speech-To-Speech Translation in the EMIME Project*
Mikko Kurimo, William Byrne, John Dines, Philip N. Garner, Matthew Gibson, Yong Guan, Teemu Hirsimäki, Reima Karhila, Simon King, Hui Liang, Keiichiro Oura, Lakshmi Saheer, Matt Shannon, Sayaki Shiota and Jilei Tian (p. 142)

(82) *Hunting for the Black Swan: Risk Mining from Text*
Jochen Leidner and Frank Schilder (p. 142)

(83) *Speech-Driven Access to the Deep Web on Mobile Devices*
Taniya Mishra and Srinivas Bangalore (p. 142)

(84) *Tools for Multilingual Grammar-Based Translation on the Web*
Aarne Ranta, Krasimir Angelov and Thomas Hallgren (p. 143)

(85) *Demonstration of a Prototype for a Conversational Companion for Reminiscing about Images*
Yorick Wilks, Roberta Catizone, Alexiei Dingli and Weiwei Cheng (p. 143)

(86) *It Makes Sense: A Wide-Coverage Word Sense Disambiguation System for Free Text*
Zhi Zhong and Hwee Tou Ng (p. 143)

16:15–16:45 ***Coffee/Tea Break, Venue A and B, Foyer***

Semantics 4

Venue A, Aula. Chair: Joyce Chai

16:45–17:10 *Learning Script Knowledge with Web Experiments*
Michaela Regneri, Alexander Koller and Manfred Pinkal (p. 143)

17:10–17:35 *Starting from Scratch in Semantic Role Labeling*
Michael Connor, Yael Gertner, Cynthia Fisher and Dan Roth (p. 144)

Dialogue

Venue A, Hall X. Chair: Adam Vogel

16:45–17:10 *Modeling Norms of Turn-Taking in Multi-Party Conversation*
Kornel Laskowski (p. 144)

17:10–17:35 *Optimising Information Presentation for Spoken Dialogue Systems*
Verena Rieser, Oliver Lemon and Xingkun Liu (p. 144)

Historical Linguistics

Venue A, Hall IX. Chair: Steven Bird

16:45–17:10 *Combining Data and Mathematical Models of Language Change*
Morgan Sonderegger and Partha Niyogi (p. 144)

17:10–17:35 *Finding Cognate Groups Using Phylogenies*
David Hall and Dan Klein (p. 145)

Decipherment

Venue B, Lecture Hall 3. Chair: Philipp Koehn

16:45–17:10 *An Exact A* Method for Deciphering Letter-Substitution Ciphers*
Eric Corlett and Gerald Penn (p. 145)

17:10–17:35 *A Statistical Model for Lost Language Decipherment*
Benjamin Snyder, Regina Barzilay and Kevin Knight (p. 145)

Tree Transducers

Venue B, Lecture Hall 4. Chair: Mark Johnson

16:45–17:10 *Efficient Inference through Cascades of Weighted Tree Transducers*
Jonathan May, Kevin Knight and Heiko Vogler (p. 145)

17:10–17:35 *A Tree Transducer Model for Synchronous Tree-Adjoining Grammars*
Andreas Maletti (p. 145)

9 Main Conference, Day 3, July 14

Overview: Main Conference - Day 3 · Wednesday, July 14

7:30–9:00	Registration. Venue A, Foyer				
9:00–10:00	Invited talk. Venue A, Aula				
10:00–10:30	*Coffee/Tea Break*				
	Venue A, Aula	Venue A, X	Venue A, IX	Venue B, 3	Venue B, 4
10:30–12:10	Parsing 3	Text Classification and Topic Models	Psycholinguistics	Semantics 5	Multimodal
12:20–13:20	Business Meeting. Venue A, Aula				
13:00–14:30	*Lunch*				
14:30–15:45	Unsupervised Parsing and Grammar Induction	Information Extraction 3	Information Retrieval	Sentiment 3	Discourse 2
15:45–16:15	*Coffee/Tea Break*				
16.15–17:30	Translation 5	Information Extraction 4	Parsing and Grammars	Word Sense Disambiguation	Generation
17:30–17:40	*Break*				
17:40–18:15 18:15-18:30	Best Paper Awards. Venue A, Aula Closing.				

Wednesday July 14, 2010

7:30–9:00 ***Registration***

Invited Talk

Venue A, Aula. Chair: Stephen Clark

9:00–10:00 ***Computational Advertising***

Andrei Broder, Yahoo! Research

Abstract: Computational advertising is an emerging new scientific sub-discipline, at the intersection of large scale search and text analysis, information retrieval, statistical modeling, machine learning, classification, optimization, and microeconomics. The central challenge of computational advertising is to find the "best match" between a given user in a given context and a suitable advertisement. The context could be a user entering a query in a search engine ("sponsored search"), a user reading a web page ("content match" and "display ads"), a user watching a movie on a portable device, and so on. The information about the user can vary from scarily detailed to practically nil. The number of potential advertisements might be in the billions. Thus, depending on the definition of "best match" this challenge leads to a variety of massive optimization and search problems, with complicated constraints. This talk will give an introduction to this area focusing on the interplay between science, engineering, and marketplace.

10:00–10:30 ***Coffee/Tea Break, Venue A, Foyer***

Parsing 3

Venue A, Aula. Chair: Jenny Rose Finkel

10:30–10:55 *Dynamic Programming for Linear-Time Incremental Parsing*
Liang Huang and Kenji Sagae (p. 147)

10:55–11:20 *Hard Constraints for Grammatical Function Labelling*
Wolfgang Seeker, Ines Rehbein, Jonas Kuhn and Josef van Genabith (p. 147)

11:20–11:45 *Simple, Accurate Parsing with an All-Fragments Grammar*
Mohit Bansal and Dan Klein (p. 147)

11:45–12:10 *Joint Syntactic and Semantic Parsing of Chinese*
Junhui Li, Guodong Zhou and Hwee Tou Ng (p. 147)

Text Classification and Topic Models

Venue A, Hall X. Chair: Diarmuid O Seaghdha

10:30–10:55 *Cross-Language Text Classification Using Structural Correspondence Learning*
Peter Prettenhofer and Benno Stein (p. 148)

10:55–11:20 *Cross-Lingual Latent Topic Extraction*
Duo Zhang, Qiaozhu Mei and ChengXiang Zhai (p. 148)

11:20–11:45 *Topic Models for Word Sense Disambiguation and Token-Based Idiom Detection*
Linlin Li, Benjamin Roth and Caroline Sporleder (p. 148)

11:45–12:10 *PCFGs, Topic Models, Adaptor Grammars and Learning Topical Collocations and the Structure of Proper Names*
Mark Johnson (p. 148)

Psycholinguistics

Venue A, Hall IX. Chair: John Hale

10:30–10:55 *A Cognitive Cost Model of Annotations Based on Eye-Tracking Data*
Katrin Tomanek, Udo Hahn, Steffen Lohmann and Jürgen Ziegler (p. 149)

10:55–11:20 *A Rational Model of Eye Movement Control in Reading*
Klinton Bicknell and Roger Levy (p. 149)

11:20–11:45 *The Influence of Discourse on Syntax: A Psycholinguistic Model of Sentence Processing*
Amit Dubey (p. 149)

11:45–12:10 *Complexity Metrics in an Incremental Right-Corner Parser*
Stephen Wu, Asaf Bachrach, Carlos Cardenas and William Schuler (p. 149)

Semantics 5

Venue B, Lecture Hall 3. Chair: Lillian Lee

10:30–10:55 **Challenge Paper:** *"Ask Not What Textual Entailment Can Do for You…"*
Mark Sammons, V.G.Vinod Vydiswaran and Dan Roth (p. 150)

10:55–11:20 *Assessing the Role of Discourse References in Entailment Inference*
Shachar Mirkin, Ido Dagan and Sebastian Pado (p. 150)

11:20–11:45 *Global Learning of Focused Entailment Graphs*
Jonathan Berant, Ido Dagan and Jacob Goldberger (p. 150)

11:45–12:10 *Modeling Semantic Relevance for Question-Answer Pairs in Web Social Communities*
Baoxun Wang, Xiaolong Wang, Chengjie Sun, Bingquan Liu and Lin Sun (p. 150)

Multimodal

Venue B, Lecture Hall 4. Chair: Alexander Koller

10:30–10:55 *How Many Words Is a Picture Worth? Automatic Caption Generation for News Images*
Yansong Feng and Mirella Lapata (p. 150)

10:55–11:20 *Generating Image Descriptions Using Dependency Relational Patterns*
Ahmet Aker and Robert Gaizauskas (p. 151)

11:20–11:45 *Incorporating Extra-Linguistic Information into Reference Resolution in Collaborative Task Dialogue*
Ryu Iida, Syumpei Kobayashi and Takenobu Tokunaga (p. 151)

11:45–12:10 *Reading between the Lines: Learning to Map High-Level Instructions to Commands*
S.R.K. Branavan, Luke Zettlemoyer and Regina Barzilay (p. 151)

12:10–12:20 ***Short Break***

ACL Business Meeting

Venue A, Aula

12:20–13:20 ***ACL Business Meeting***

13:00–14:30 ***Lunch (Complimentary)***

Unsupervised Parsing and Grammar Induction

Venue A, Aula. Chair: Yusuke Miyao

14:30–14:55 *Profiting from Mark-Up: Hyper-Text Annotations for Guided Parsing*
Valentin I. Spitkovsky, Daniel Jurafsky and Hiyan Alshawi (p. 151)

14:55–15:20 *Phylogenetic Grammar Induction*
Taylor Berg-Kirkpatrick and Dan Klein (p. 152)

15:20–15:45 *Improved Unsupervised POS Induction through Prototype Discovery*
Omri Abend, Roi Reichart and Ari Rappoport (p. 152)

Information Extraction 3

Venue A, Hall X. Chair: James R. Curran

14:30–14:55 *Extraction and Approximation of Numerical Attributes from the Web*
Dmitry Davidov and Ari Rappoport (p. 152)

14:55–15:20 *Learning Word-Class Lattices for Definition and Hypernym Extraction*
Roberto Navigli and Paola Velardi (p. 152)

15:20–15:45 *On Learning Subtypes of the Part-Whole Relation: Do Not Mix Your Seeds*
Ashwin Ittoo and Gosse Bouma (p. 153)

Information Retrieval

Venue A, Hall IX. Chair: Christof Monz

14:30–14:55 *Understanding the Semantic Structure of Noun Phrase Queries*
Xiao Li (p. 153)

14:55–15:20 *Multilingual Pseudo-Relevance Feedback: Performance Study of Assisting Languages*
Manoj Kumar Chinnakotla, Karthik Raman and Pushpak Bhattacharyya (p. 153)

15:20–15:45 *Wikipedia as Sense Inventory to Improve Diversity in Web Search Results*
Celina Santamaría, Julio Gonzalo and Javier Artiles (p. 153)

Sentiment 3

Venue B, Lecture Hall 3. Chair: Dragomir Radev

14:30–14:55 *A Unified Graph Model for Sentence-Based Opinion Retrieval*
Binyang Li, Lanjun Zhou, Shi Feng and Kam-Fai Wong (p. 154)

14:55–15:20 *Generating Fine-Grained Reviews of Songs from Album Reviews*
Swati Tata and Barbara Di Eugenio (p. 154)

15:20–15:45 *A Study of Information Retrieval Weighting Schemes for Sentiment Analysis*
Georgios Paltoglou and Mike Thelwall (p. 154)

Discourse 2

Venue B, Lecture Hall 4. Chair: Jian Su

14:30–14:55 **Survey Paper:** *Supervised Noun Phrase Coreference Research: The First Fifteen Years*
Vincent Ng (p. 154)

14:55–15:20 *Unsupervised Event Coreference Resolution with Rich Linguistic Features*
Cosmin Bejan and Sanda Harabagiu (p. 155)

15:20–15:45 *Coreference Resolution across Corpora: Languages, Coding Schemes, and Preprocessing Information*
Marta Recasens and Eduard Hovy (p. 155)

15:45–16:15 ***Coffee/Tea Break, Venue A and B, Foyer***

Translation 5

Venue A, Aula. Chair: Min Zhang

16:15–16:40 *Constituency to Dependency Translation with Forests*
Haitao Mi and Qun Liu (p. 155)

16:40–17:05 *Learning to Translate with Source and Target Syntax*
David Chiang (p. 155)

17:05–17:30 *Discriminative Modeling of Extraction Sets for Machine Translation*
John DeNero and Dan Klein (p. 155)

Information Extraction 4

Venue A, Hall X. Chair: Massimo Poesio

16:15–16:40 *Detecting Experiences from Weblogs*
Keun Chan Park, Yoonjae Jeong and Sung Hyon Myaeng (p. 156)

16:40–17:05 *Experiments in Graph-Based Semi-Supervised Learning Methods for Class-Instance Acquisition*
Partha Pratim Talukdar and Fernando Pereira (p. 156)

17:05–17:30 *Learning Arguments and Supertypes of Semantic Relations Using Recursive Patterns*
Zornitsa Kozareva and Eduard Hovy (p. 156)

Parsing and Grammars

Venue A, Hall IX. Chair: David Weir

16:15–16:40 *A Transition-Based Parser for 2-Planar Dependency Structures*
Carlos Gómez-Rodríguez and Joakim Nivre (p. 156)

16:40–17:05 *Viterbi Training for PCFGs: Hardness Results and Competitiveness of Uniform Initialization*
Shay Cohen and Noah A. Smith (p. 157)

17:05–17:30 *A Generalized-Zero-Preserving Method for Compact Encoding of Concept Lattices*
Matthew Skala, Victoria Krakovna, János Kramár and Gerald Penn (p. 157)

Word Sense Disambiguation

Venue B, Lecture Hall 3. Chair: Sebastian Pado

16:15–16:40 *Knowledge-Rich Word Sense Disambiguation Rivaling Supervised Systems*
Simone Paolo Ponzetto and Roberto Navigli (p. 157)

16:40–17:05 *All Words Domain Adapted WSD: Finding a Middle Ground between Supervision and Unsupervision*
Mitesh Khapra, Anup Kulkarni, Saurabh Sohoney and Pushpak Bhattacharyya (p. 157)

17:05–17:30 *Combining Orthogonal Monolingual and Multilingual Sources of Evidence for All Words WSD*
Weiwei Guo and Mona Diab (p. 158)

Generation

Venue B, Lecture Hall 4. Chair: Johanna D. Moore

16:15–16:40 *Phrase-Based Statistical Language Generation Using Graphical Models and Active Learning*
Francois Mairesse, Milica Gasic, Filip Jurcicek, Simon Keizer, Blaise Thomson, Kai Yu and Steve Young (p. 158)

16:40–17:05 *Plot Induction and Evolutionary Search for Story Generation*
Neil McIntyre and Mirella Lapata (p. 158)

17:05–17:30 *Automated Planning for Situated Natural Language Generation*
Konstantina Garoufi and Alexander Koller (p. 158)

17:30–17:40 ***Short Break***

Best Paper Awards

Venue A, Aula. Chair: Stephen Clark

17:40–17:50 ***Best Paper Awards Ceremony***

17:50–18:15 **Best Long Paper:** *Beyond NomBank: A Study of Implicit Arguments for Nominal Predicates*
Matthew Gerber and Joyce Chai (p. 159)

Closing of ACL 2010 Main Conference

Venue A, Aula

18:15–18:30 ***Closing***

10 CoNLL-2010, July 15–16

The Fourteenth Conference on Computational Natural Language Learning

Venue A, Aula.

Conference co-chairs: Anoop Sarkar and Mirella Lapata

Thursday, July 15, 2010

9:00–9:15 ***Opening Remarks***

Session 1: Parsing

9:15–9:40 *Improvements in Unsupervised Co-Occurrence-Based Parsing*
Christian Hänig (p. 160)

9:40–10:05 *Viterbi Training Improves Unsupervised Dependency Parsing*
Valentin I. Spitkovsky, Hiyan Alshawi, Daniel Jurafsky and Christopher D. Manning (p. 160)

10:05–10:30 *Driving Semantic Parsing from the World's Response*
James Clarke, Dan Goldwasser, Ming-Wei Chang and Dan Roth (p. 160)

10:30–11:00 ***Break***

Session 2: Grammar Induction

11:00–11:25 *Efficient, Correct, Unsupervised Learning for Context-Sensitive Languages*
Alexander Clark (p. 161)

11:25–11:50 *Identifying Patterns for Unsupervised Grammar Induction*
Jesús Santamaría and Lourdes Araujo (p. 161)

11:50–12:15 *Learning Better Monolingual Models with Unannotated Bilingual Text*
David Burkett, Slav Petrov, John Blitzer and Dan Klein (p. 161)

12:15–14:15 ***Lunch***

Invited Talk

14:15–15:30 *Clueless: Explorations in Unsupervised, Knowledge-Lean Extraction of Lexical-Semantic Information*
Lillian Lee (p. 161)

15:30–16:00 ***Break***

Shared Task Session 1: Overview and Oral Presentations

16:00–16:20 *The CoNLL 2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text*
Richárd Farkas, Veronika Vincze, György Móra, János Csirik and György Szarvas (p. 162)

16:20–16:30 *A Cascade Method for Detecting Hedges and their Scope in Natural Language Text*
Buzhou Tang, Xiaolong Wang, Xuan Wang, Bo Yuan and Shixi Fan (p. 162)

16:30–16:40 *Detecting Speculative Language using Syntactic Dependencies and Logistic Regression*
Andreas Vlachos and Mark Craven (p. 162)

16:40–16:50 *A Hedgehop over a Max-margin Framework using Hedge Cues*
Maria Georgescul (p. 163)

16:50–17:00 *Detecting Hedge Cues and their Scopes with Average Perceptron*
Feng Ji, Xipeng Qiu and Xuanjing Huang (p. 163)

17:00–17:10 *Memory-based Resolution of In-sentence Scopes of Hedge Cues*
Roser Morante, Vincent Van Asch and Walter Daelemans (p. 163)

17:10–17:20 *Resolving Speculation: MaxEnt Cue Classification and Dependency-Based Scope Rules*
Erik Velldal, Lilja Øvrelid and Stephan Oepen (p. 163)

17:20–17:30 *Combining Manual Rules and Supervised Learning for Hedge Cue and Scope Detection*
Marek Rei and Ted Briscoe (p. 163)

17:30–18:00 ***Shared Task Discussion Panel***

Friday, July 16, 2010

Invited Talk

9:15–10:30 *Bayesian Hidden Markov Models and Extensions*
Zoubin Ghahramani (p. 165)

10:30–11:00 ***Break***

Joint Poster Session: Main Conference and Shared Task Posters

11:00–12:30 ***Poster Presentations***

Posters: Main Conference

(21) *Improved Unsupervised POS Induction Using Intrinsic Clustering Quality and a Zipfian Constraint*
Roi Reichart, Raanan Fattal and Ari Rappoport (p. 165)

(22) *Syntactic and Semantic Structure for Opinion Expression Detection*
Richard Johansson and Alessandro Moschitti (p. 165)

16:25–16:50 *On Reverse Feature Engineering of Syntactic Tree Kernels*
Daniele Pighin and Alessandro Moschitti (p. 172)

16:50–17:15 *Inspecting the Structural Biases of Dependency Parsing Algorithms*
Yoav Goldberg and Michael Elhadad (p. 173)

17:15–17:45 ***SIGNLL Business Meeting and Best Paper Award***

11 Workshops, July 15–16

WS1: SemEval-2010: 5th International Workshop on Semantic Evaluations
July 15–16. Venue B, Lecture Hall 4.
Chairs: Katrin Erk and Carlo Strapparava.

WS2: Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR
July 15–16. Venue B, Lecture Hall 3.
Chairs: Chris Callison-Burch, Philipp Koehn, Christof Monz and Kay Peterson.

WS3: The Fourth Linguistic Annotation Workshop (The LAW IV)
July 15–16. Venue A, Hall X.
Chairs: Nianwen Xue and Massimo Poesio.

WS4: 2010 Workshop on Biomedical Natural Language Processing (BioNLP 2010)
July 15. Venue A, Hall IX.
Chairs: K. Bretonnel Cohen, Dina Demner-Fushman, Sophia Ananiadou, John Pestian, Jun'ichi Tsujii and Bonnie Webber.

WS5: Cognitive Modeling and Computational Linguistics
July 15. Venue A, Room VIII.
Chair: John Hale.

WS6: NLP and Linguistics: Finding the Common Ground
July 16. Venue A, Hall IX.
Chairs: Fei Xia, William Lewis and Lori Levin.

WS7: 11th Meeting of ACL Special Interest Group in Computational Morphology and Phonology (SIGMORPHON)
July 15. Venue A, Room XI.
Chairs: Jeffrey Heinz, Lynne Cahill and Richard Wicentowski.

WS8: TextGraphs-5: Graph-based Methods for Natural Language Processing
July 16. Venue A, Hall IV.
Chairs: Carmen Banea, Alessandro Moschitti, Swapna Somasundaran and Fabio Massimo Zanzotto.

WS9: Named Entities Workshop (NEWS 2010)
July 16. Venue A, Room VIII.
Chairs: A Kumaran and Haizhou Li.

WS10: Applications of Tree Automata in Natural Language Processing
July 16. Venue A, Room XI.
Chairs: Frank Drewes and Marco Kuhlmann.

WS11: Domain Adaptation for Natural Language Processing (DANLP)
July 15. Venue A, Hall IV.
Chairs: Hal Daumé III, Tejaswini Deoskar, David McClosky, Barbara Plank and Jörg Tiedemann.

WS12: Companionable Dialogue Systems
July 15. Venue A, Room II.
Chairs: Yorick Wilks, Morena Danieli and Björn Gambäck.

WS13: GEMS-2010 Geometric Models of Natural Language Semantics
July 16. Venue A, Room II.
Chairs: Roberto Basili and Marco Pennacchiotti.

WS1: SemEval-2010: 5th International Workshop on Semantic Evaluations

July 15–16. Venue B, Lecture Hall 4.

Chairs: Katrin Erk and Carlo Strapparava

Thursday, July 15, 2010

Task description papers

9:00–9:20 *SemEval-2010 Task 1: Coreference Resolution in Multiple Languages*
Marta Recasens, Lluís Màrquez, Emili Sapena, M. Antònia Martí, Mariona Taulé, Véronique Hoste, Massimo Poesio and Yannick Versley

9:20–9:40 *SemEval-2010 Task 2: Cross-Lingual Lexical Substitution*
Rada Mihalcea, Ravi Sinha and Diana McCarthy

9:40–10:00 *SemEval-2010 Task 3: Cross-Lingual Word Sense Disambiguation*
Els Lefever and Véronique Hoste

10:00–10:20 *SemEval-2010 Task 5 : Automatic Keyphrase Extraction from Scientific Articles*
Su Nam Kim, Olena Medelyan, Min-Yen Kan and Timothy Baldwin

10:20–10:40 *SemEval-2010 Task 7: Argument Selection and Coercion*
James Pustejovsky, Anna Rumshisky, Alex Plotnick, Elisabetta Jezek, Olga Batiukova and Valeria Quochi

10:40–11:00 ***Break***

Task description papers

11:00–11:20 *SemEval-2010 Task 8: Multi-Way Classification of Semantic Relations Between Pairs of Nominals*
Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Sebastian Pado, Marco Pennacchiotti, Lorenza Romano and Stan Szpakowicz

11:20–11:40 *SemEval-2 Task 9: The Interpretation of Noun Compounds Using Paraphrasing Verbs and Prepositions*
Cristina Butnariu, Su Nam Kim, Preslav Nakov, Diarmuid Ó Séaghdha, Stan Szpakowicz and Tony Veale

11:40–12:00 *SemEval-2010 Task 10: Linking Events and Their Participants in Discourse*
Josef Ruppenhofer, Caroline Sporleder, Roser Morante, Collin Baker and Martha Palmer

12:00–12:20 *SemEval-2010 Task 12: Parser Evaluation using Textual Entailments*
Deniz Yuret, Aydin Han and Zehra Turgut

12:20–12:40 *SemEval-2010 Task 13: TempEval-2*
Marc Verhagen, Roser Sauri, Tommaso Caselli and James Pustejovsky

12:40–14:00 ***Lunch***

Task description papers

14:00–14:20 *SemEval-2010 Task 14: Word Sense Induction Disambiguation*
Suresh Manandhar, Ioannis Klapaftis, Dmitriy Dligach and Sameer Pradhan

14:20–14:40 *SemEval-2010 Task: Japanese WSD*
Manabu Okumura, Kiyoaki Shirai, Kanako Komiya and Hikaru Yokono

14:40–15:00 *SemEval-2010 Task 17: All-words Word Sense Disambiguation on a Specific Domain*
Eneko Agirre, Oier Lopez de Lacalle, Christiane Fellbaum, Shu-Kai Hsieh, Maurizio Tesconi, Monica Monachini, Piek Vossen and Roxanne Segers

15:00–15:20 *SemEval-2010 Task 18: Disambiguating Sentiment Ambiguous Adjectives*
Yunfang Wu and Peng Jin

15:20–16:00 ***Break***

Poster Session

16:00–17:30 ***Poster Presentations***

(101) *RelaxCor: A Global Relaxation Labeling Approach to Coreference Resolution*
Emili Sapena, Lluís Padró and Jordi Turmo

(102) *SUCRE: A Modular System for Coreference Resolution*
Hamidreza Kobdani and Hinrich Schütze

(103) *UBIU: A Language-Independent System for Coreference Resolution*
Desislava Zhekova and Sandra Kübler

(104) *Corry: a System for Coreference Resolution*
Olga Uryupina

(105) *BART: A Multilingual Anaphora Resolution System*
Samuel Broscheit, Massimo Poesio, Simone Paolo Ponzetto, Kepa Joseba Rodriguez, Lorenza Romano, Olga Uryupina, Yannick Versley and Roberto Zanoli

(106) *TANL-1: Coreference Resolution by Parse Analysis and Similarity Clustering*
Giuseppe Attardi, Maria Simi and Stefano Dei Rossi

(107) *FCC: Modeling Probabilities with GIZA++ for Task #2 and #3 of SemEval-2*
Darnes Vilariño Ayala, Carlos Balderas Posada, David Eduardo Pinto Avendaño, Miguel Rodríguez Hernández and Saul León Silverio

(108) *Combining Dictionaries and Contextual Information for Cross-Lingual Lexical Substitution*
Wilker Aziz and Lucia Specia

(125) *KP-Miner: Participation in SemEval-2*
Samhaa R. El-Beltagy and Ahmed Rafea

(126) *UvT: The UvT Term Extraction System in the Keyphrase Extraction task*
Kalliopi Zervanou

(127) *UNITN: Part-Of-Speech Counting in Relation Extraction*
Fabio Celli

(128) *FBK_NK: a WordNet-based System for Multi-Way Classification of Semantic Relations*
Matteo Negri and Milen Kouylekov

(129) *JU: A Supervised Approach to Identify Semantic Relations from Paired Nominals*
Santanu Pal, Partha Pakray, Dipankar Das and Sivaji Bandyopadhyay

(130) *TUD: semantic relatedness for relation classification*
György Szarvas and Iryna Gurevych

(131) *FBK-IRST: Semantic Relation Extraction using Cyc*
Kateryna Tymoshenko and Claudio Giuliano

(132) *ISTI@SemEval-2 Task #8: Boosting-Based Multiway Relation Classification*
Andrea Esuli, Diego Marcheggiani and Fabrizio Sebastiani

(133) *ISI: Automatic Classification of Relations Between Nominals Using a Maximum Entropy Classifier*
Stephen Tratz and Eduard Hovy

(134) *ECNU: Effective Semantic Relations Classification without Complicated Features or Multiple External Corpora*
Yuan Chen, Man Lan, Jian Su, Zhi Min Zhou and Yu Xu

(135) *UCD-Goggle: A Hybrid System for Noun Compound Paraphrasing*
Guofu Li, Alejandra Lopez-Fernandez and Tony Veale

(136) *UCD-PN: Selecting General Paraphrases Using Conditional Probability*
Paul Nulty and Fintan Costello

Friday, July 16, 2010

System papers

9:00–9:15 *COLEPL and COLSLM: An Unsupervised WSD Approach to Multilingual Lexical Substitution, Tasks 2 and 3 SemEval 2010*
Weiwei Guo and Mona Diab

9:15–9:30 *UBA: Using Automatic Translation and Wikipedia for Cross-Lingual Lexical Substitution*
Pierpaolo Basile and Giovanni Semeraro

9:30–9:45 *HUMB: Automatic Key Term Extraction from Scientific Articles in GROBID*
Patrice Lopez and Laurent Romary

9:45–10:00 *UTDMet: Combining WordNet and Corpus Data for Argument Coercion Detection*
Kirk Roberts and Sanda Harabagiu

10:00–10:15 *UTD: Classifying Semantic Relations by Combining Lexical and Semantic Resources*
Bryan Rink and Sanda Harabagiu

10:15–10:30 *UvT: Memory-based pairwise ranking of paraphrasing verbs*
Sander Wubben

10:30–11:00 ***Break***

System papers

11:00–11:15 *SEMAFOR: Frame Argument Resolution with Log-Linear Models*
Desai Chen, Nathan Schneider, Dipanjan Das and Noah A. Smith

11:15–11:30 *Cambridge: Parser Evaluation using Textual Entailment by Grammatical Relation Comparison*
Laura Rimell and Stephen Clark

11:30–11:45 *MARS: A Specialized RTE System for Parser Evaluation*
Rui Wang and Yi Zhang

11:45–12:00 *TRIPS and TRIOS System for TempEval-2: Extracting Temporal Information from Text*
Naushad UzZaman and James Allen

12:00–12:15 *TIPSem (English and Spanish): Evaluating CRFs and Semantic Roles in TempEval-2*
Hector Llorens, Estela Saquete Boro and Borja Navarro

12:15–12:30 *CityU-DAC: Disambiguating Sentiment-Ambiguous Adjectives within Context*
Bin Lu and Benjamin K. Tsou

12:30–14:00 ***Lunch***

14:00–15:30 ***Panel***

15:30–16:00 ***Break***

Posters Session

16:00–17:30 ***Poster Presentations***

(101) *VENSES++: Adapting a deep semantic processing system to the identification of null instantiations*
Sara Tonelli and Rodolfo Delmonte

(102) *CLR: Linking Events and Their Participants in Discourse Using a Comprehensive FrameNet Dictionary*
Ken Litkowski

(103) *PKU_HIT: An Event Detection System Based on Instances Expansion and Rich Syntactic Features*
Shiqi Li, Peng-Yuan Liu, Tiejun Zhao, Qin Lu and Hanjing Li

WS2: Joint Fifth Workshop on Statistical Machine Translation and MetricsMATR

July 15–16. Venue B, Lecture Hall 3.

Chairs: Chris Callison-Burch, Philipp Koehn, Christof Monz and Kay Peterson

Thursday, July 15, 2010

8:45–9:00 ***Opening Remarks***

Full Paper Session 1

9:00–9:25 *A Semi-supervised Word Alignment Algorithm with Partial Manual Alignments*
Qin Gao, Nguyen Bach and Stephan Vogel

9:25–9:50 *Fast Consensus Hypothesis Regeneration for Machine Translation*
Boxing Chen, George Foster and Roland Kuhn

Shared Translation Task

9:50–10:15 *Findings of the 2010 Joint Workshop on Statistical Machine Translation and Metrics for Machine Translation*
Chris Callison-Burch, Philipp Koehn, Christof Monz, Kay Peterson, Mark Przybocki and Omar Zaidan

10:15–10:45 ***Boaster Session 1: Translation Task***

10:45–11:00 ***Morning Break***

Poster Session: Translation Task

11:00–12:30 ***Poster Presentations***

(101) *LIMSI's Statistical Translation Systems for WMT'10*
Alexandre Allauzen, Josep M. Crego, İlknur Durgar El-Kahlout and François Yvon

(102) *2010 Failures in English-Czech Phrase-Based MT*
Ondřej Bojar and Kamil Kos

(103) *An Empirical Study on Development Set Selection Strategy for Machine Translation Learning*
Hui Cong, Zhao Hai, Lu Bao-Liang and Song Yan

(104) *The University of Maryland Statistical Machine Translation System for the Fifth Workshop on Machine Translation*
Vladimir Eidelman, Chris Dyer and Philip Resnik

(105) *Further Experiments with Shallow Hybrid MT Systems*
Christian Federmann, Andreas Eisele, Yu Chen, Sabine Hunsicker, Jia Xu and Hans Uszkoreit

(106) *Improved Features and Grammar Selection for Syntax-Based MT*
Greg Hanneman, Jonathan Clark and Alon Lavie

(123) *UPV-PRHLT English–Spanish System for WMT10*
Germán Sanchis-Trilles, Jesús Andrés-Ferrer, Guillem Gascó, Jesús González Rubio, Pascual Martínez-Gómez, Martha-Alicia Rocha, Joan-Andreu Sánchez and Francisco Casacuberta

(124) *Reproducible Results in Parsing-Based Machine Translation: The JHU Shared Task Submission*
Lane Schwartz

(125) *Vs and OOVs: Two Problems for Translation between German and English*
Sara Stymne, Maria Holmqvist and Lars Ahrenberg

(126) *To Cache or not to Cache? Experiments with Adaptive Models in Statistical Machine Translation*
Jörg Tiedemann

(127) *Applying Morphological Decompositions to Statistical Machine Translation*
Sami Virpioja, Jaakko Väyrynen, Andre Mansikkaniemi and Mikko Kurimo

(128) *Maximum Entropy Translation Model in Dependency-Based MT Framework*
Zdeněk Žabokrtský, Martin Popel and David Mareček

(129) *UCH-UPV English–Spanish system for WMT10*
Francisco Zamora-Martinez and Germán Sanchis-Trilles

(130) *Hierarchical Phrase-Based MT at the Charles University for the WMT 2010 Shared Task*
Daniel Zeman

12:30–14:00 ***Lunch***

Invited Talk

14:00–15:00 ***Invited Talk***
Hermann Ney

Full Paper Session 2

15:05–15:30 *Incremental Decoding for Phrase-based Statistical Machine Translation*
Baskaran Sankaran, Ajeet Grewal and Anoop Sarkar

15:30–16:00 ***Afternoon Break***

Full Paper Session 3

16:00–16:25 *How to Avoid Burning Ducks: Combining Linguistic Analysis and Corpus Statistics for German Compound Processing*
Fabienne Fritzinger and Alexander Fraser

16:25–16:50 *Chunk-based Verb Reordering in VSO Sentences for Arabic-English Statistical Machine Translation*
Arianna Bisazza and Marcello Federico

16:50–17:15 *Head Finalization: A Simple Reordering Rule for SOV Languages*
Hideki Isozaki, Katsuhito Sudoh, Hajime Tsukada and Kevin Duh

17:15–17:40 *Aiding Pronoun Translation with Co-Reference Resolution*
Ronan Le Nagard and Philipp Koehn

Friday, July 16, 2010

Shared Task Presentations

9:00–10:00 ***Overview: MetricsMATR***
10:00–10:30 ***Discussion***
10:30–10:45 ***Boaster Session***
10:45–11:00 ***Morning Break***

Poster Sessions

11:00–12:30 ***Poster Presentations***

Poster: Full Paper

(101) *Jane: Open Source Hierarchical Translation, Extended with Reordering and Lexicon Models*
David Vilar, Daniel Stein, Matthias Huck and Hermann Ney

Posters: System Combination Task

(102) *MANY: Open Source MT System Combination at WMT'10*
Loïc Barrault

(103) *Adaptive Model Weighting and Transductive Regression for Predicting Best System Combinations*
Ergun Bicici and S. Serdar Kozat

(104) *L1 Regularized Regression for Reranking and System Combination in Machine Translation*
Ergun Bicici and Deniz Yuret

(105) *An Augmented Three-Pass System Combination Framework: DCU Combination System for WMT 2010*
Jinhua Du, Pavel Pecina and Andy Way

(106) *The UPV-PRHLT Combination System for WMT 2010*
Jesús González Rubio, Germán Sanchis-Trilles, Joan-Andreu Sánchez, Jesús Andrés-Ferrer, Guillem Gascó, Pascual Martínez-Gómez, Martha-Alicia Rocha and Francisco Casacuberta

(107) *CMU Multi-Engine Machine Translation for WMT 2010*
Kenneth Heafield and Alon Lavie

(108) *CMU System Combination via Hypothesis Selection for WMT'10*
Almut Silja Hildebrand and Stephan Vogel

(109) *JHU System Combination Scheme for WMT 2010*
Sushant Narsale

(110) *The RWTH System Combination System for WMT 2010*
Gregor Leusch and Hermann Ney

(111) *BBN System Description for WMT10 System Combination Task*
Antti-Veikko Rosti, Bing Zhang, Spyros Matsoukas and Richard Schwartz

Posters: Metrics Task

(112) *LRscore for Evaluating Lexical and Reordering Quality in MT*
Alexandra Birch and Miles Osborne

(113) *Document-level Automatic MT Evaluation based on Discourse Representations*
Elisabet Comelles, Jesus Gimenez, Lluis Marquez, Irene Castellon and Victoria Arranz

(114) *METEOR-NEXT and the METEOR Paraphrase Tables: Improved Evaluation Support for Five Target Languages*
Michael Denkowski and Alon Lavie

(115) *Normalized Compression Distance Based Measures for MetricsMATR 2010*
Marcus Dobrinkat, Tero Tapiovaara, Jaakko Väyrynen and Kimmo Kettunen

(116) *The DCU Dependency-Based Metric in WMT-MetricsMATR 2010*
Yifan He, Jinhua Du, Andy Way and Josef van Genabith

(117) *TESLA: Translation Evaluation of Sentences with Linear-programming-based Analysis*
Chang Liu, Daniel Dahlmeier and Hwee Tou Ng

(118) *The Parameter-optimized ATEC Metric for MT Evaluation*
Billy Wong and Chunyu Kit

12:30–14:00 ***Lunch***

Full Paper Session 4

14:00–14:25 *A Unified Approach to Minimum Risk Training and Decoding*
Abhishek Arun, Barry Haddow and Philipp Koehn

14:25–14:50 *N-best Reranking by Multitask Learning*
Kevin Duh, Katsuhito Sudoh, Hajime Tsukada, Hideki Isozaki and Masaaki Nagata

14:50–15:15 *Taming Structured Perceptrons on Wild Feature Vectors*
Ralf Brown

15:15–15:40 *Translation Model Adaptation by Resampling*
Kashif Shah, Loïc Barrault and Holger Schwenk

15:40–16:00 ***Afternoon Break***

Full Paper Session 5

16:00–16:25 *Integration of Multiple Bilingually-Learned Segmentation Schemes into Statistical Machine Translation*
Michael Paul, Andrew Finch and Eiichiro Sumita

16:25–16:50 *Improved Translation with Source Syntax Labels*
Hieu Hoang and Philipp Koehn

16:50–17:15 *Divide and Translate: Improving Long Distance Reordering in Statistical Machine Translation*
Katsuhito Sudoh, Kevin Duh, Hajime Tsukada, Tsutomu Hirao and Masaaki Nagata

17:15–17:40 *Decision Trees for Lexical Smoothing in Statistical Machine Translation*
Rabih Zbib, Spyros Matsoukas, Richard Schwartz and John Makhoul

WS3: The Fourth Linguistic Annotation Workshop (The LAW IV)

July 15–16. Venue A, Hall X.

Chairs: Nianwen Xue and Massimo Poesio

Thursday, July 15, 2010

08:40–08:50 ***Opening Remarks***

Session I

Chair: Nianwen Xue

08:50–09:15 *EmotiBlog: a Finer-Grained and More Precise Learning of Subjectivity Expression Models*
Ester Boldrini, Alexandra Balahur, Patricio Martínez-Barco and Andrés Montoyo Guijarro

09:15–09:40 *Error-tagged Learner Corpus of Czech*
Jirka Hana, Alexandr Rosen, Svatava Škodová and Barbora Štindlová

09:40–10:05 *Annotation Scheme for Social Network Extraction from Text*
Apoorv Agarwal, Owen Rambow and Rebecca Passonneau

10:05–10:30 *Agile Corpus Annotation in Practice: An Overview of Manual and Automatic Annotation of CVs*
Beatrice Alex, Claire Grover, Rongzhou Shen and Mijail Kabadjov

10:30–11:00 ***Break***

Session II

Chair: Martha Palmer

11:00–11:25 *Consistency Checking for Treebank Alignment*
Markus Dickinson and Yvonne Samuelsson

11:25–11:50 *Anveshan: A Framework for Analysis of Multiple Annotators' Labeling Behavior*
Vikas Bhardwaj, Rebecca Passonneau, Ansaf Salleb-Aouissi and Nancy Ide

11:50–12:15 *Influence of Pre-annotation on POS-tagged Corpus Development*
Karën Fort and Benoît Sagot

12:15–12:40 *To Annotate More Accurately or to Annotate More*
Dmitriy Dligach, Rodney Nielsen and Martha Palmer

12:40–13:50 ***Lunch***

Session III

Chair: Manfred Stede

13:50–14:15 *Annotating Underquantification*
Aurelie Herbelot and Ann Copestake

14:15–14:40 *PropBank Annotation of Multilingual Light Verb Constructions*
Jena D. Hwang, Archna Bhatia, Claire Bonial, Aous Mansouri, Ashwini Vaidya, Nianwen Xue and Martha Palmer

14:40–15:05 *Retrieving Correct Semantic Boundaries in Dependency Structure*
Jinho Choi and Martha Palmer

15:05–15:30 *Complex Predicates Annotation in a Corpus of Portuguese*
Iris Hendrickx, Amália Mendes, Sílvia Pereira, Anabela Gonçalves and Inês Duarte

15:30–16:00 ***Break***

Poster session

Chair: Nianwen Xue

16:00–17:30 ***Poster Presentations***

(1) *Using an Online Tool for the Documentation of Edo Language*
Ota Ogie

(2) *Cross-Lingual Validity of PropBank in the Manual Annotation of French*
Lonneke van der Plas, Tanja Samardzic and Paola Merlo

(3) *Characteristics of High Agreement Affect Snnotation in Text*
Cecilia Ovesdotter Alm

(4) *The Deep Re-annotation in a Chinese Scientific Treebank*
Kun Yu, Xiangli Wang, Yusuke Miyao, Takuya Matsuzaki and Jun'ichi Tsujii

(5) *The Unified Annotation of Syntax and Discourse in the Copenhagen Dependency Treebanks*
Matthias Buch-Kromann and Iørn Korzen

(6) *Identifying Sources of Inter-Annotator Variation: Evaluating Two Models of Argument Analysis*
Barbara White

(7) *Dependency-Based PropBanking of Clinical Finnish*
Katri Haverinen, Filip Ginter, Timo Viljanen, Veronika Laippala and Tapio Salakoski

(8) *Building the Syntactic Reference Corpus of Medieval French Using NotaBene RDF Annotation Tool*
Nicolas Mazziotta

(9) *Chunking German: An Unsolved Problem*
Sandra Kübler, Kathrin Beck, Erhard Hinrichs and Heike Telljohann

(10) *Proposal for MWE Annotation in Running Text*
Iris Hendrickx, Amália Mendes and Sandra Antunes

(11) *A Feature Type Classification for Therapeutic Purposes: a preliminary evaluation with non-expert speakers*
Gianluca E. Lebani and Emanuele Pianta

(12) *Annotating Korean Demonstratives*
Sun-Hee Lee and Jae-young Song

(21) *Creating and Exploiting a Resource of Parallel Parses*
Christian Chiarcos, Kerstin Eckart and Julia Ritz

(22) *From Descriptive Annotation to Grammar Specification*
Lars Hellan

(23) *An Annotation Schema for Preposition Senses in German*
Antje Müller, Olaf Hülscher, Claudia Roch, Katja Kesselmeier, Tobias Stadtfeld, Jan Strunk and Tibor Kiss

(24) *OTTO: A Transcription and Management Tool for Historical Texts*
Stefanie Dipper, Lara Kresse, Martin Schnurrenberger and Seong-Eun Cho

(25) *Multimodal Annotation of Conversational Data*
Philippe Blache, Roxane Bertrand, Emmanuel Bruno, Brigitte Bigi, Robert Espesser, Gaelle Ferre, Mathilde Guardiola, Daniel Hirst, Ning Tan, Edlira Cela, Jean-Claude Martin, Stéphane Rauzy, Mary-Annick Morel, Elisabeth Murisasco and Irina Nesterenko

(26) *Combining Parallel Treebanks and Geo-Tagging*
Martin Volk, Anne Goehring and Torsten Marek

(27) *Challenges of Cheap Resource Creation*
Jirka Hana and Anna Feldman

(28) *Discourse Relation Configurations in Turkish and an Annotation Environment*
Berfin Aktaş, Cem Bozşahin and Deniz Zeyrek

(29) *An Overview of the CRAFT Concept Annotation Guidelines*
Michael Bada, Miriam Eckert, Martha Palmer and Lawrence Hunter

(30) *Syntactic Tree Queries in Prolog*
Gerlof Bouma

(31) *An Integrated Tool for Annotating Historical Corpora*
Pablo Picasso Feliciano de Faria, Fabio Natanael Kepler and Maria Clara Paixão de Sousa

(32) *The Revised Arabic PropBank*
Wajdi Zaghouani, Mona Diab, Aous Mansouri, Sameer Pradhan and Martha Palmer

Friday, July 16, 2010

Session IV

Chair: Massimo Poesio

08:50–09:15 *PackPlay: Mining Semantic Data in Collaborative Games*
Nathan Green, Paul Breimyer, Vinay Kumar and Nagiza Samatova

09:15–09:40 *A Proposal for a Configurable Silver Standard*
Udo Hahn, Katrin Tomanek, Elena Beisswanger and Erik Faessler

09:40–10:05 *A Hybrid Model for Annotating Named Entity Training Corpora*
Robert Voyer, Valerie Nygaard, Will Fitzgerald and Hannah Copperman

10:05–10:30 *Anatomy of Annotation Schemes: Mapping to GrAF*
Nancy Ide and Harry Bunt

10:30–11:00 ***Break***

Session V

Chair: Nancy Ide

11:00–11:25 *Annotating Participant Reference in English Spoken Conversation*
John Niekrasz and Johanna D. Moore

11:25–11:50 *Design and Evaluation of Shared Prosodic Annotation for Spontaneous French Speech: From Expert Knowledge to Non-Expert Annotation*
Anne Lacheret-Dujour, Nicolas Obin and Mathieu Avanzi

11:50–12:15 *Depends on What the French Say - Spoken Corpus Annotation With and Beyond Syntactic Functions*
José Deulofeu, Lucie Duffort, Kim Gerdes, Sylvain Kahane and Paola Pietrandrea

12:15–12:40 *The Annotation Scheme of the Turkish Discourse Bank and An Evaluation of Inconsistent Annotations*
Deniz Zeyrek, Işin Demirşahin, Ayişiği Sevdik-Çalli, Hale Ögel Balaban, Ihsan Yalçinkaya and Ümit Deniz Turan

12:40–13:00 ***Closing remarks***

WS4: 2010 Workshop on Biomedical Natural Language Processing (BioNLP 2010)

July 15. Venue A, Hall IX.

Chairs: K. Bretonnel Cohen, Dina Demner-Fushman, Sophia Ananiadou, John Pestian, Jun'ichi Tsujii and Bonnie Webber

9:00–9:15 ***Opening Remarks***

Session 1: Extraction

9:15–9:40 *Two Strong Baselines for the BioNLP 2009 Event Extraction Task*
Andreas Vlachos

9:40–10:05 *Recognizing Biomedical Named Entities using Skip-chain Conditional Random Fields*
Jingchen Liu, Minlie Huang and Xiaoyan Zhu

10:05–10:30 *Event Extraction for Post-Translational Modifications*
Tomoko Ohta, Sampo Pyysalo, Makoto Miwa, Jin-Dong Kim and Jun'ichi Tsujii

10:30–11:00 ***Morning Coffee Break***

Session 2

11:00–12:00 ***Keynote Speaker***
Text Mining and Intelligence
W. John Wilbur

12:05–12:30 *Scaling up Biomedical Event Extraction to the Entire PubMed*
Jari Björne, Filip Ginter, Sampo Pyysalo, Jun'ichi Tsujii and Tapio Salakoski

12:30–14:00 ***Lunch Break***

Session 3: Foundations

14:00–14:25 *A Comparative Study of Syntactic Parsers for Event Extraction*
Makoto Miwa, Sampo Pyysalo, Tadayoshi Hara and Jun'ichi Tsujii

14:25–14:50 *Arguments of Nominals in Semantic Interpretation of Biomedical Text*
Halil Kilicoglu, Marcelo Fiszman, Graciela Rosemblat, Sean Marimpietri and Thomas Rindflesch

Session 4: High-Level Tasks

14:50–15:15 *Improving Summarization of Biomedical Documents using Word Sense Disambiguation*
Laura Plaza, Mark Stevenson and Alberto Díaz

15:30–16:00 ***Afternoon Coffee Break***

Session 4: High-Level Tasks, continued

16:00–16:25 *Cancer Stage Prediction Based on Patient Online Discourse*
Mukund Jha and Noemie Elhadad

16:25–16:50 *An Exploration of Mining Gene Expression Mentions and their Anatomical Locations from Biomedical Text*
Martin Gerner, Goran Nenadic and Casey M. Bergman

16:50–17:00 ***Poster Boaster Session and Conclusions***

17:00–17:30 ***Poster Presentations***

(37) *Exploring Surface-level Heuristics for Negation and Speculation Discovery in Clinical Texts*
Emilia Apostolova and Noriko Tomuro

(38) *Disease Mention Recognition with Specific Features*
Md. Faisal Mahbub Chowdhury and Alberto Lavelli

(39) *Extraction of Disease-Treatment Semantic Relations from Biomedical Sentences*
Oana Frunza and Diana Inkpen

(40) *Identifying the Information Structure of Scientific Abstracts: An Investigation of Three Different Schemes*
Yufan Guo, Anna Korhonen, Maria Liakata, Ilona Silins, Lin Sun and Ulla Stenius

(41) *Reconstruction of Semantic Relationships from Their Projections in Biomolecular Domain*
Juho Heimonen, Jari Björne and Tapio Salakoski

(42) *Towards Internet-Age Pharmacovigilance: Extracting Adverse Drug Reactions from User Posts in Health-Related Social Networks*
Robert Leaman, Laura Wojtulewicz, Ryan Sullivan, Annie Skariah, Jian Yang and Graciela Gonzalez

(43) *Semantic Role Labeling of Gene Regulation Events: Preliminary Results*
Roser Morante

(44) *Ontology-Based Extraction and Summarization of Protein Mutation Impact Information*
Nona Naderi and René Witte

(45) *Extracting Distinctive Features of Swine (H1N1) Flu through Data Mining Clinical Documents*
Heekyong Park and Jinwook Choi

(46) *Towards Event Extraction from Full Texts on Infectious Diseases*
Sampo Pyysalo, Tomoko Ohta, Han-Cheol Cho, Dan Sullivan, Chunhong Mao, Bruno Sobral, Sophia Ananiadou and Jun'ichi Tsujii

(47) *Applying the TARSQI Toolkit to Augment Text Mining of EHRs*
Amber Stubbs and Benjamin Harshfield

(48) *Integration of Static Relations to Enhance Event Extraction from Text*
Sofie Van Landeghem, Sampo Pyysalo, Tomoko Ohta and Yves Van de Peer

WS5: Cognitive Modeling and Computational Linguistics

July 15. Venue A, Room VIII.

Chair: John Hale

Language change at multiple levels

9:00–9:30 *Using Sentence Type Information for Syntactic Category Acquisition*
Stella Frank, Sharon Goldwater and Frank Keller

9:30–10:00 *Did Social Networks Shape Language Evolution? A Multi-Agent Cognitive Simulation*
David Reitter and Christian Lebiere

10:00–10:30 *Syntactic Adaptation in Language Comprehension*
Alex Fine, Ting Qian, T. Florian Jaeger and Robert Jacobs

10:30–11:00 ***Morning Break***

Parsing and memory

11:00–11:30 *HHMM Parsing with Limited Parallelism*
Tim Miller and William Schuler

11:30–12:00 *The Role of Memory in Superiority Violation Gradience*
Marisa Ferrara Boston

12:00–14:00 ***Lunch Break***

Corpus-based modeling

14:00–14:30 *Close = Relevant? The Role of Context in Efficient Language Production*
Ting Qian and T. Florian Jaeger

14:30–15:00 *Predicting Cognitively Salient Modifiers of the Constitutive Parts of Concepts*
Gerhard Kremer and Marco Baroni

15:00–15:30 *Towards a Data-Driven Model of Eye Movement Control in Reading*
Mattias Nilsson and Joakim Nivre

15:30–16:00 ***Afternoon Break***

Information-theoretical approaches

16:00–16:30 *Modeling the Noun Phrase versus Sentence Coordination Ambiguity in Dutch: Evidence from Surprisal Theory*
Harm Brouwer, Hartmut Fitz and John Hoeks

16:30–17:00 *Uncertainty Reduction as a Measure of Cognitive Processing Effort*
Stefan Frank

WS6: NLP and Linguistics: Finding the Common Ground

July 16. Venue A, Hall IX.

Chairs: Fei Xia, William Lewis and Lori Levin

8:45–8:50 ***Opening Remarks***

8:50–9:50 ***Invited Talk***
The Human Language Project: Uniting Computational Linguistics with Documentary Linguistics
Steven Bird

Paper Session 1

9:50–10:10 *Modeling and Encoding Traditional Wordlists for Machine Applications*
Shakthi Poornima and Jeff Good

10:10–10:30 *Evidentiality for Text Trustworthiness Detection*
Su Qi, Huang Chu-Ren and Chen Kai-yun

10:30–11:00 ***Morning Break***

Panel Session 1: NLP helps Linguistics

11:00–12:00 ***Presentations***
Presentation and discussion from panelists
Hal Daumé III, Alexis Dimitriadis, Erhard Hinrichs and Dipti Misra Sharma
On the Role of NLP in Linguistics
Dipti Misra Sharma
Matching Needs and Resources: How NLP Can Help Theoretical Linguistics
Alexis Dimitriadis

Paper Session 2

12:00–12:20 *Grammar-Driven versus Data-Driven: Which Parsing System is More Affected by Domain Shifts?*
Barbara Plank and Gertjan van Noord

12:20–12:40 *A Cross-Lingual Induction Technique for German Adverbial Participles*
Sina Zarrieß, Aoife Cahill, Jonas Kuhn and Christian Rohrer

12:40–14:10 ***Lunch***

Paper Session 3

14:10–14:30 *You Talking to Me? A Predictive Model for Zero Auxiliary Constructions*
Andrew Caines and Paula Buttery

14:30–14:50 *Cross-Lingual Variation of Light Verb Constructions: Using Parallel Corpora and Automatic Alignment for Linguistic Research*
Tanja Samardzic and Paola Merlo

14:50–15:10 *No Sentence is Too Confusing to Ignore*
Paul Cook and Suzanne Stevenson

15:10–15:30 *Consonant Co-occurrence in Stems Across Languages: Automatic Analysis and Visualization of a Phonotactic Constraint*
Thomas Mayer, Christian Rohrdantz, Frans Plank, Peter Bak, Miriam Butt and Daniel A. Keim

15:30–16:00 ***Afternoon Break***

Panel Session 2: Linguistics helps NLP

16:00–17:00 ***Presentations***

Presentation and discussion from panelists
Julia Hockenmaier, Eduard Hovy and Owen Rambow

Injecting Linguistics into NLP through Annotation
Eduard Hovy

17:00–17:30 ***Group discussion and closing***

WS7: 11th Meeting of ACL Special Interest Group in Computational Morphology and Phonology (SIGMORPHON)

July 15. Venue A, Room XI.

Chairs: Jeffrey Heinz, Lynne Cahill and Richard Wicentowski

Session

9:00–9:30 *Instance-Based Acquisition of Vowel Harmony*
Fred Mailhot

9:30–10:00 *Verifying Vowel Harmony Typologies*
Sara Finley

10:00–10:30 *Complexity of the Acquisition of Phonotactics in Optimality Theory*
Giorgio Magri

10:30–11:00 ***Morning Break***

Session

11:00–11:30 *Maximum Likelihood Estimation of Feature-Based Distributions*
Jeffrey Heinz and Cesar Koirala

11:30–12:00 *A Method for Compiling Two-Level Rules with Multiple Contexts*
Kimmo Koskenniemi and Miikka Silfverberg

12:00–12:30 *Exploring Dialect Phonetic Variation Using PARAFAC*
Jelena Prokic and Tim Van de Cruys

12:30–14:00 ***Lunch***

Session

14:00–14:30 *Quantitative Evaluation of Competing Syllable Parses*
Jason A. Shaw and Adamantios I. Gafos

14:30–15:00 *Toward a Totally Unsupervised, Language-Independent Method for the Syllabification of Written Texts*
Thomas Mayer

15:00–15:30 *Comparing Canonicalizations of Historical German Text*
Bryan Jurish

15:30–16:00 ***Afternoon Break***

Session

16:00–16:30 *Semi-Supervised Learning of Concatenative Morphology*
Oskar Kohonen, Sami Virpioja and Krista Lagus

16:30–17:00 *Morpho Challenge 2005-2010: Evaluations and Results*
Mikko Kurimo, Sami Virpioja, Ville Turunen and Krista Lagus

17.00 ***Business Meeting***

WS8: TextGraphs-5: Graph-based Methods for Natural Language Processing

July 16. Venue A, Hall IV.

Chairs: Carmen Banea, Alessandro Moschitti, Swapna Somasundaran and Fabio Massimo Zanzotto

09:00–09:10 ***Welcome to TextGraphs-5***

Session 1: Lexical Clustering and Disambiguation

09:10–09:30 *Graph-based Clustering for Computational Linguistics: a Survey*
Zheng Chen and Heng Ji

09:30–09:50 *Towards the Automatic Creation of a Wordnet from a Term-based Lexical Network*
Hugo Gonçalo Oliveira and Paulo Gomes

09:50–10:10 *An Investigation on the Influence of Frequency on the Lexical Organization of Verbs*
Daniel German, Aline Villavicencio and Maity Siqueira

10:10–10:30 *Robust and Efficient Page Rank for Word Sense Disambiguation*
Diego De Cao, Roberto Basili, Matteo Luciani, Francesco Mesiano and Riccardo Rossi

10:30–11:00 ***Coffee Break***

Session 2: Clustering Languages and Dialects

11:00–11:20 *Hierarchical Spectral Partitioning of Bipartite Graphs to Cluster Dialects and Identify Distinguishing Features*
Martijn Wieling and John Nerbonne

11:20–11:40 *A Character-Based Intersection Graph Approach to Linguistic Phylogeny*
Jessica Enright

Invited Talk

11:40–12:40 *Spectral Approaches to Learning in the Graph Domain*
Edwin Hancock

12:50–13:50 ***Lunch Break***

Session 3: Lexical Similarity and Its application

13:50–14:10 *Cross-lingual Comparison between Distributionally Determined Word Similarity Networks*
Olof Görnerup and Jussi Karlgren

14:10–14:30 *Co-occurrence Cluster Features for Lexical Substitutions in Context*
Chris Biemann

14:30–14:50 *Contextually-Mediated Semantic Similarity Graphs for Topic Segmentation*
Geetu Ambwani and Anthony Davis

14:50–15:10 *MuLLinG: MultiLevel Linguistic Graphs for Knowledge Extraction*
Vincent Archer

15:10–15:30 *Experiments with CST-based Multidocument Summarization*
Maria Lucia Castro Jorge and Thiago Pardo

15:30–16:00 ***Coffee Break***

Special Session on Opinion Mining

16:00–16:20 *Distinguishing between Positive and Negative Opinions with Complex Network Features*
Diego Raphael Amancio, Renato Fabbri, Osvaldo Novais Oliveira Jr., Maria das Graças Volpe Nunes and Luciano da Fontoura Costa

16:20–16:40 *Image and Collateral Text in Support of Auto-annotation and Sentiment Analysis*
Pamela Zontone, Giulia Boato, Jonathon Hare, Paul Lewis, Stefan Siersdorfer and Enrico Minack

16:40–17:00 *Aggregating Opinions: Explorations into Graphs and Media Content Analysis*
Gabriele Tatzl and Christoph Waldhauser

Session 5: Spectral Approaches

17:00–17:20 *Eliminating Redundancy by Spectral Relaxation for Multi-Document Summarization*
Fumiyo Fukumoto, Akina Sakai and Yoshimi Suzuki

17:20–17:40 *Computing Word Senses by Semantic Mirroring and Spectral Graph Partitioning*
Martin Fagerlund, Magnus Merkel, Lars Eldén and Lars Ahrenberg

17:40–18:00 ***Final Wrap-up***

WS9: Named Entities Workshop (NEWS 2010)

July 16. Venue A, Room VIII.

Chairs: A Kumaran and Haizhou Li

Session 1: Oral

9:00–9:15 ***Opening Remarks***
A Kumaran and Haizhou Li

9:15–10:00 ***Keynote Speech***
Dan Roth

10:00–10:30 *Transliteration Generation and Mining with Limited Training Resources*
Sittichai Jiampojamarn, Kenneth Dwyer, Shane Bergsma, Aditya Bhargava, Qing Dou, Mi-Young Kim and Grzegorz Kondrak

10:30–11:00 ***Morning Break***

Session 2: Oral

11:00–11:20 *Transliteration Using a Phrase-Based Statistical Machine Translation System to Re-Score the Output of a Joint Multigram Model*
Andrew Finch and Eiichiro Sumita

11:20–11:40 *Transliteration Mining with Phonetic Conflation and Iterative Training*
Kareem Darwish

Session 3: Poster Presentation

11:40–12:40 ***Poster Presentations***

(37) *Language Independent Transliteration Mining System Using Finite State Automata Framework*
Sara Noeman and Amgad Madkour

(38) *Reranking with Multiple Features for Better Transliteration*
Yan Song, Chunyu Kit and Hai Zhao

(39) *Syllable-Based Thai-English Machine Transliteration*
Chai Wutiwiwatchai and Ausdang Thangthai

(40) *English to Indian Languages Machine Transliteration System at NEWS 2010*
Amitava Das, Tanik Saikh, Tapabrata Mondal, Asif Ekbal and Sivaji Bandyopadhyay

(41) *Mining Transliterations from Wikipedia Using Pair HMMs*
Peter Nabende

(42) *Phrase-Based Transliteration with Simple Heuristics*
Avinesh PVS and Ankur Parikh

12:40–14:00 ***Lunch Break***

Session 4: Oral

14:00–14:20 *Classifying Wikipedia Articles into NE's Using SVM's with Threshold Adjustment*
Iman Saleh, Kareem Darwish and Aly Fahmy

14:20–14:40 *Assessing the Challenge of Fine-Grained Named Entity Recognition and Classification*
Asif Ekbal, Eva Sourjikova, Anette Frank and Simone Paolo Ponzetto

14:40–15:00 *Using Deep Belief Nets for Chinese Named Entity Categorization*
Yu Chen, You Ouyang, Wenjie Li, Dequan Zheng and Tiejun Zhao

15:00–15:20 *Simplified Feature Set for Arabic Named Entity Recognition*
Ahmed Abdul Hamid and Kareem Darwish

15:20–16:00 ***Break***

Session 5: Oral

16:00–16:20 *Think Globally, Apply Locally: Using Distributional Characteristics for Hindi Named Entity Identification*
Shalini Gupta and Pushpak Bhattacharyya

16:20–16:40 *Rule-Based Named Entity Recognition in Urdu*
Kashif Riaz

16:40–17:00 *CONE: Metrics for Automatic Evaluation of Named Entity Co-Reference Resolution*
Bo Lin, Rushin Shah, Robert Frederking and Anatole Gershman

17:00–17:10 ***Closing***

WS10: Applications of Tree Automata in Natural Language Processing

July 16. Venue A, Room XI.

Chairs: Frank Drewes and Marco Kuhlmann

09:00–09:15 ***Opening Remarks***

09:15–10:30 ***Invited Talk***
Kevin Knight

10:30–11:00 ***Coffee Break***

Full Paper Session 1

11:00–11:30 *Preservation of Recognizability for Synchronous Tree Substitution Grammars*
Zoltán Fülöp, Andreas Maletti and Heiko Vogler

11:30–12:00 *A Decoder for Probabilistic Synchronous Tree Insertion Grammars*
Steve DeNeefe, Kevin Knight and Heiko Vogler

12:00–12:30 *Parsing and Translation Algorithms Based on Weighted Extended Tree Transducers*
Andreas Maletti and Giorgio Satta

12:30–14:00 ***Lunch Break***

Full Paper Session 2

14:00–14:30 *Millstream Systems – a Formal Model for Linking Language Modules by Interfaces*
Suna Bensch and Frank Drewes

14:30–15:00 *Transforming Lexica as Trees*
Mark-Jan Nederhof

15:00–15:30 *n-Best Parsing Revisited*
Matthias Büchse, Daniel Geisler, Torsten Stüber and Heiko Vogler

15:30–16:00 ***Coffee Break***

Quickfire Presentations

16:00–16:15 *Tree Automata Techniques and the Learning of Semantic Grammars*
Michael Minock

16:15–16:30 *Do We Really Want a Single Tree to Cover the Whole Sentence?*
Aravind Joshi

16:30–16:45 *The Tree Automata Workbench 'Marbles'*
Frank Drewes

16:45–17:00 *Requirements on a Tree Transformation Model for Machine Translation*
Andreas Maletti

17:00–17:30 ***Discussion***

WS11: Domain Adaptation for Natural Language Processing (DANLP)

July 15. Venue A, Hall IV.

Chairs: Hal Daumé III, Tejaswini Deoskar, David McClosky, Barbara Plank and Jörg Tiedemann

9:15–9:30 ***Opening***
Barbara Plank

9:30–10:30 ***Invited Talk***
Semi-supervised Domain Adaptation: From Practice to Theory
John Blitzer

10:30–11:00 ***Morning Break***

Session I

11:00–11:25 *Adaptive Parameters for Entity Recognition with Perceptron HMMs*
Massimiliano Ciaramita and Olivier Chapelle

11:30–11:55 *Context Adaptation in Statistical Machine Translation Using Models with Exponentially Decaying Cache*
Jörg Tiedemann

12:00–12:25 *Domain Adaptation to Summarize Human Conversations*
Oana Sandu, Giuseppe Carenini, Gabriel Murray and Raymond Ng

12:30–14:00 ***Lunch***

Session II

14:00–14:25 *Exploring Representation-Learning Approaches to Domain Adaptation*
Fei Huang and Alexander Yates

14:30–14:55 *Using Domain Similarity for Performance Estimation*
Vincent Van Asch and Walter Daelemans

15:00–15:25 *Self-Training without Reranking for Parser Domain Adaptation and Its Impact on Semantic Role Labeling*
Kenji Sagae

15:30–16:00 ***Afternoon Break***

Session III

16:00–16:25 *Domain Adaptation with Unlabeled Data for Dialog Act Tagging*
Anna Margolis, Karen Livescu and Mari Ostendorf

16:30–16:55 *Frustratingly Easy Semi-Supervised Domain Adaptation*
Hal Daumé III, Abhishek Kumar and Avishek Saha

17:00–17:45 ***Panel Discussion***
John Blitzer, Walter Daelemans, Hal Daumé III, Jing Jiang and Khalil Sima'an

WS12: Companionable Dialogue Systems

July 15. Venue A, Room II.

Chairs: Yorick Wilks, Morena Danieli and Björn Gambäck

Invited Paper Session

09:00–9:15 ***Welcome***

09:15–10:30 *Do's and Don'ts for Software Companions*
David Traum

10:30–11:00 ***Morning Break***

Session

11:00–11:30 *Episodic Memory for Companion Dialogue*
Gregor Sieber and Brigitte Krenn

11:30–12:00 *MANA for the Ageing*
David M W Powers, Martin H Luerssen, Trent W Lewis, Richard E Leibbrandt, Marissa Milne, John Pashalis and Kenneth Treharne

12:00–12:30 *Is a Companion a Distinctive Kind of Relationship with a Machine?*
Yorick Wilks

12:30–14:00 ***Lunch Break***

Session

14:00–14:30 *"Hello Emily, How are You Today?" - Personalised Dialogue in a Toy to Engage Children.*
Carole Adam, Lawrence Cavedon and Lin Padgham

14:30–15:00 *A Robot in the Kitchen*
Peter Wallis

15:00–15:30 *An Embodied Dialogue System with Personality and Emotions*
Stasinos Konstantopoulos

15:30–16:00 ***Afternoon Break***

Session

16:00–16:30 *How was Your Day?*
Stephen Pulman, Johan Boye, Marc Cavazza, Cameron Smith and Raúl Santos de la Cámara

16:30–17:00 *VCA: An Experiment With A Multiparty Virtual Chat Agent*
Samira Shaikh, Tomek Strzalkowski, Sarah Taylor and Nick Webb

17:00–17:30 ***Wrap up discussion of the day's issues***

WS13: GEMS-2010 Geometric Models of Natural Language Semantics

July 16. Venue A, Room II.

Chairs: Roberto Basili and Marco Pennacchiotti

9:25–9:30 ***Welcome and Opening***

Session: Geometry and Semantics

9:30–10:00 *Capturing Nonlinear Structure in Word Spaces Through Dimensionality Reduction*
David Jurgens and Keith Stevens

10:00–10:30 *Manifold Learning for the Semi-Supervised Induction of FrameNet Predicates: an Empirical Investigation*
Danilo Croce and Daniele Previtali

10:30–11:00 ***Coffee Break***

Invited Talk

11:00–12:10 *What is Word Meaning, Really? (And How Can Distributional Models Help Us Describe It?)*
Katrin Erk

Session: Lexical Acquisition 1

12:10–12:40 *Relatedness Curves for Acquiring Paraphrases*
Georgiana Dinu and Grzegorz Chrupała

12:40–13:10 *A Regression Model of Adjective-Noun Compositionality in Distributional Semantics*
Emiliano Guevara and Daniele Previtali

13:10–14:30 ***Lunch Break***

Session: Lexical Acquisition 2

14:30–15:00 *Semantic Composition with Quotient Algebras*
Daoud Clarke, Rudi Lutz and David Weir

15:00–15.30 *Expectation Vectors: A Semiotics Inspired Approach to Geometric Lexical-Semantic Representation*
Justin Washtell

15:30–16:00 ***Coffee Break***

Session: Computational Aspects

16:00–16:30 *Sketch Techniques for scaling Distributional Similarity to the Web*
Amit Goyal, Jagadeesh Jagaralamudi, Hal Daumé III and Suresh Venkatasubramanian

16:30–17:00 *Active Learning for Constrained Dirichlet Process Mixture Models*
Andreas Vlachos, Zoubin Ghahramani and Ted Briscoe

17:00–17:55 ***Panel***

17:55–18:00 ***Closing Remarks***

12 ACL 2010 Main Conference Abstracts

ACL 2010 Main Conference Abstracts: Monday, July 12

Parsing 1, 10:30–11:45, Venue A, Aula

Efficient Third-Order Dependency Parsers
Terry Koo and Michael Collins
We present algorithms for higher-order dependency parsing that are "third-order" in the sense that they can evaluate sub-structures containing three dependencies, and "efficient" in the sense that they require only $O(n^4)$ time. Importantly, our new parsers can utilize both sibling-style and grandchild-style interactions. We evaluate our parsers on the Penn Treebank and Prague Dependency Treebank, achieving unlabeled attachment scores of 93.04% and 87.38%, respectively.

Dependency Parsing and Projection Based on Word-Pair Classification
Wenbin Jiang and Qun Liu
In this paper we describe an intuitionistic method for dependency parsing, where a classifier is used to determine whether a pair of words forms a dependency edge. And we also propose an effective strategy for dependency projection, where the dependency relationships of the word pairs in the source language are projected to the word pairs of the target language, leading to a set of classification instances rather than a complete tree. Experiments show that, the classifier trained on the projected classification instances significantly outperforms previous projected dependency parsers. More importantly, when this classifier is integrated into a maximum spanning tree (MST) dependency parser, obvious improvement is obtained over the MST baseline.

Bitext Dependency Parsing with Bilingual Subtree Constraints
Wenliang Chen, Jun'ichi Kazama and Kentaro Torisawa
This paper proposes a dependency parsing method that uses bilingual constraints to improve the accuracy of parsing bilingual texts (bitexts). In our method, a target-side tree fragment that corresponds to a source-side tree fragment is identified via word alignment and mapping rules that are automatically learned. Then it is verified by checking the subtree list that is collected from large scale automatically parsed data on the target side. Our method, thus, requires gold standard trees only on the source side of a bilingual corpus in the training phase, unlike the joint parsing model, which requires gold standard trees on the both sides. Compared to the reordering constraint model, which requires the same training data as ours, our method achieved higher accuracy because of richer bilingual constraints. Experiments on the translated portion of the Chinese Treebank show that our system outperforms monolingual parsers by 2.93 points for Chinese and 1.64 points for English.

Semantics 1, 10:30–11:45, Venue A, Hall X

Computing Weakest Readings
Alexander Koller and Stefan Thater
We present an efficient algorithm for computing the weakest readings of semantically ambiguous sentences. A corpus-based evaluation with a large-scale grammar shows that our algorithm reduces

over 80% of sentences to one or two readings, in negligible runtime, and thus makes it possible to work with semantic representations derived by deep large-scale grammars.

Identifying Generic Noun Phrases
Nils Reiter and Anette Frank
This paper presents a supervised approach for identifying generic noun phrases in context. Generic statements express rule-like knowledge about kinds or events. Therefore, their identification is important for the automatic construction of knowledge bases. In particular, the distinction between generic and non-generic statements is crucial for the correct encoding of generic and instance-level information. Generic expressions have been studied extensively in formal semantics. Building on this work, we explore a corpus-based learning approach for identifying generic NPs, using selections of linguistically motivated features. Our results perform well above the baseline and existing prior work.

Structural Semantic Relatedness: A Knowledge-Based Method to Named Entity Disambiguation
Xianpei Han and Jun Zhao
Name ambiguity problem has raised urgent demands for efficient, high-quality named entity disambiguation methods. In recent years, the increasing availability of large-scale, rich semantic knowledge sources (such as Wikipedia and WordNet) creates new opportunities to enhance the named entity disambiguation by developing algorithms which can exploit these knowledge sources at best. The problem is that these knowledge sources are heterogeneous and most of the semantic knowledge within them is embedded in complex structures, such as graphs and networks. This paper proposes a knowledge-based method, called Structural Semantic Relatedness (SSR), which can enhance the named entity disambiguation by capturing and leveraging the structural semantic knowledge in multiple know-ledge sources. Empirical results show that, in comparison with the classical BOW based methods and social network based methods, our method can significantly improve the disambiguation performance by respectively 8.7% and 14.7%.

Spoken Language, 10:30–11:45, Venue A, Hall IX

Correcting Errors in Speech Recognition with Articulatory Dynamics
Frank Rudzicz
We introduce a novel mechanism for incorporating articulatory dynamics into speech recognition with the theory of task dynamics. This system reranks sentence-level hypotheses by the likelihoods of their hypothetical articulatory realizations which are derived from relationships learned with aligned acoustic/articulatory data. Experiments compare this with two baseline systems, namely an acoustic hidden Markov model and a dynamic Bayes network augmented with discretized representations of the vocal tract. Our system based on task dynamics reduces word-error rates significantly by 10.2% relative to the best baseline models.

Learning to Adapt to Unknown Users: Referring Expression Generation in Spoken Dialogue Systems
Srinivasan Janarthanam and Oliver Lemon
We present a data-driven approach to learn user-adaptive referring expression generation (REG) policies for spoken dialogue systems. Referring expressions can be difficult to understand in technical domains where users may not know the technical 'jargon' names of the domain entities. In such cases, dialogue systems must be able to model the user's (lexical) domain knowledge and use appropriate referring expressions. We present a reinforcement learning (RL) framework in which the system learns REG policies which can adapt to unknown users online. Furthermore, unlike supervised learning methods which require a large corpus of expert adaptive behaviour to train on, we show that effective adaptive policies can be learned from a small dialogue corpus of non-adaptive human-machine interaction, by using a RL framework and a statistical user

simulation. We show that in comparison to adaptive hand-coded baseline policies, the learned policy performs significantly better, with an 18.6% average increase in adaptation accuracy. The best learned policy also takes less dialogue time (average 1.07 min less) than the best hand-coded policy. This is because the learned policies can adapt online to changing evidence about the user's domain expertise.

A Risk Minimization Framework for Extractive Speech Summarization
Shih-Hsiang Lin and Berlin Chen
In this paper, we formulate extractive summarization as a risk minimization problem and propose a unified probabilistic framework that naturally combines supervised and unsupervised summarization models to inherit their individual merits as well as to overcome their inherent limitations. In addition, the introduction of various loss functions also provides the sum-marization framework with a flexible but systematic way to render the redundancy and coherence relationships among sentences and between sentences and the whole document, respectively. Experiments on speech summarization show that the methods deduced from our framework are very competitive with existing summarization approaches.

Resources and Evaluation, 10:30–11:45, Venue B, Lecture Hall 3

The Human Language Project: Building a Universal Corpus of the World's Languages
Steven Abney and Steven Bird
We present a grand challenge to build a corpus that will include all of the world's languages, in a consistent structure that permits large-scale cross-linguistic processing, enabling the study of universal linguistics. The focal data types, bilingual texts and lexicons, relate each language to one of a set of reference languages. We propose that the ability to train systems to translate into and out of a given language be the yardstick for determining when we have successfully captured a language. We call on the computational linguistics community to begin work on this Universal Corpus, pursuing the many strands of activity described here, as their contribution to the global effort to document the world's linguistic heritage before more languages fall silent.

Bilingual Lexicon Generation Using Non-Aligned Signatures
Daphna Shezaf and Ari Rappoport
Bilingual lexicons are fundamental resources. Modern automated lexicon generation methods usually require parallel corpora, which are not available for most language pairs. Lexicons can be generated using non-parallel corpora or a pivot language, but such lexicons are noisy. We present an algorithm for generating a high quality lexicon from a noisy one, which only requires an independent corpus for each language. Our algorithm introduces non-aligned signatures (NAS), a cross-lingual word context similarity score that avoids the over-constrained and inefficient nature of alignment-based methods. We use NAS to eliminate incorrect translations from the generated lexicon. We evaluate our method by improving the quality of noisy Spanish-Hebrew lexicons generated from two pivot English lexicons. Our algorithm substantially outperforms other lexicon generation methods.

Automatic Evaluation Method for Machine Translation Using Noun-Phrase Chunking
Hiroshi Echizen-ya and Kenji Araki
As described in this paper, we propose a new automatic evaluation method for machine translation using noun-phrase chunking. Our method correctly determines the matching words between two sentences using corresponding noun phrases. Moreover, our method determines the similarity between two sentences in terms of the noun-phrase order of appearance. Evaluation experiments were conducted to calculate the correlation among human judgments, along with the scores produced using automatic evaluation methods for MT outputs obtained from the 12 machine translation systems in NTCIR-7. Experimental results show that our method obtained

the highest correlations among the methods in both sentence-level adequacy and fluency.

Information Extraction 1, 10:30–11:45, Venue B, Lecture Hall 4

Open Information Extraction Using Wikipedia
Fei Wu and Daniel S. Weld
Information-extraction (IE) systems seek to distill semantic relations from natural language text, but most systems use supervised learning of relation-specific examples and are thus limited by the availability of training data. Open IE systems such as TextRunner, on the other hand, aim to handle the unbounded number of relations found on the Web. But how well can these open systems perform? This paper presents WOE, an open IE system which improves dramatically on TextRunner's precision and recall. The key to WOE's performance is a novel form of self-supervised learning for open extractors — using heuristic matches between Wikipedia infobox attribute values and corresponding sentences to construct training data. Like TextRunner, WOE's extractor eschews lexicalized features and handles an unbounded set of semantic relations. WOE can operate in two modes: when restricted to POS tag features, it runs as quickly as TextRunner, but when set to use dependency-parse features its precision and recall rise even higher.

SystemT: An Algebraic Approach to Declarative Information Extraction
Laura Chiticariu, Rajasekar Krishnamurthy, Yunyao Li, Sriram Raghavan, Frederick Reiss and Shivakumar Vaithyanathan
As information extraction (IE) becomes more central to enterprise applications, rule-based IE engines have become increasingly important. In this paper, we describe SystemT, a rule-based IE system whose basic design removes the expressivity and performance limitations of current systems based on cascading grammars. SystemT uses a declarative rule language, AQL, and an optimizer that generates high-performance algebraic execution plans for AQL rules. We compare SystemT's approach against cascading grammars, both theoretically and with a thorough experimental evaluation. Our results show that SystemT can deliver result quality comparable to the state-of-the-art and an order of magnitude higher annotation throughput.

Extracting Social Networks from Literary Fiction
David Elson, Nicholas Dames and Kathleen McKeown
We present a method for extracting social networks from literature, namely, nineteenth-century British novels and serials. We derive the networks from dialogue interactions, and thus our method depends on the ability to determine when two characters are in conversation. Our approach involves character name chunking, quoted speech attribution and conversation detection given the set of quotes. We extract features from the social networks and examine their correlation with one another, as well as with metadata such as the novel's setting. Our results provide evidence that the majority of novels in this time period do not fit two characterizations provided by literacy scholars. Instead, our results suggest an alternative explanation for differences in social networks.

Short Talks: Translation 1, 11:55–13:15, Venue A, Aula

Pseudo-Word for Phrase-Based Machine Translation
Xiangyu Duan, Min Zhang and Haizhou Li
The pipeline of most Phrase-Based Statistical Machine Translation (PB-SMT) systems starts from automatically word aligned parallel corpus. But word appears to be too fine-grained in some cases such as non-compositional phrasal equivalences, where no clear word alignments exist. Using words as inputs to PB-SMT pipeline has inborn deficiency. This paper proposes pseudo-word as a new start point for PB-SMT pipeline. Pseudo-word is a kind of basic multi-word expression that characterizes minimal sequence of consecutive words in sense of translation. By casting pseudo-word searching problem into a parsing framework, we search for pseudo-words in a monolingual

way and a bilingual synchronous way. Experiments show that pseudo-word significantly outperforms word for PB-SMT model in both travel translation domain and news translation domain.

Hierarchical Search for Word Alignment

Jason Riesa and Daniel Marcu

We present a simple yet powerful hierarchical search algorithm for automatic word alignment. Our algorithm induces a forest of alignments from which we can efficiently extract a ranked k-best list. We score a given alignment within the forest with a flexible, linear discriminative model incorporating hundreds of features, and trained on a relatively small amount of annotated data. We report results on Arabic-English word alignment and translation tasks. Our model outperforms a GIZA++ Model-4 baseline by 6.3 points in F-measure, yielding a 1.1 BLEU score increase over a state-of-the-art syntax-based machine translation system.

Paraphrase Lattice for Statistical Machine Translation

Takashi Onishi, Masao Utiyama and Eiichiro Sumita

Lattice decoding in statistical machine translation (SMT) is useful in speech translation and in the translation of German because it can handle input ambiguities such as speech recognition ambiguities and German word segmentation ambiguities. We show that lattice decoding is also useful for handling input variations. Given an input sentence, we build a lattice which represents paraphrases of the input sentence. We call this a paraphrase lattice. Then, we give the paraphrase lattice as an input to the lattice decoder. The decoder selects the best path for decoding. Using these paraphrase lattices as inputs, we obtained significant gains in BLEU scores for IWSLT and Europarl datasets.

A Joint Rule Selection Model for Hierarchical Phrase-Based Translation

Lei Cui, Dongdong Zhang, Mu Li, Ming Zhou and Tiejun Zhao

In hierarchical phrase-based SMT systems, statistical models are integrated to guide the hierarchical rule selection for better translation performance. Previous work mainly focused on the selection of either the source side of a hierarchical rule or the target side of a hierarchical rule rather than considering both of them simultaneously. This paper presents a joint model to predict the selection of hierarchical rules. The proposed model is estimated based on four sub-models where the rich context knowledge from both source and target sides is leveraged. Our method can be easily incorporated into the practical SMT systems with the log-linear model framework. The experimental results show that our method can yield significant improvements in performance.

Learning Lexicalized Reordering Models from Reordering Graphs

Jinsong Su, Yang Liu, Yajuan Lv, Haitao Mi and Qun Liu

Lexicalized reordering models play a crucial role in phrase-based translation systems. They are usually learned from the word-aligned bilingual corpus by examining the reordering relations of adjacent phrases. Instead of just checking whether there is one phrase adjacent to a given phrase, we argue that it is important to take the number of adjacent phrases into account for better estimations of reordering models. We propose to use a structure named reordering graph, which represents all phrase segmentations of a sentence pair, to learn lexicalized reordering models efficiently. Experimental results on the NIST Chinese-English test sets show that our approach significantly outperforms the baseline method.

Filtering Syntactic Constraints for Statistical Machine Translation

Hailong Cao and Eiichiro Sumita

Source language parse trees offer very useful but imperfect reordering constraints for statistical machine translation. A lot of effort has been made for soft applications of syntactic constraints. We alternatively propose the selective use of syntactic constraints. A classifier is built automatically to decide whether a node in the parse trees should be used as a reordering constraint or not. Using this information yields a 0.8 BLEU point improvement over a full constraint-based system.

Diversify and Combine: Improving Word Alignment for Machine Translation on Low-Resource Languages
Bing Xiang, Yonggang Deng and Bowen Zhou
We present a novel method to improve word alignment quality and eventually the translation performance by producing and combining complementary word alignments for low-resource languages. Instead of focusing on the improvement of a single set of word alignments, we generate multiple sets of diversified alignments based on different motivations, such as linguistic knowledge, morphology and heuristics. We demonstrate this approach on an English-to-Pashto translation task by combining the alignments obtained from syntactic reordering, stemming, and partial words. The combined alignment outperforms the baseline alignment, with significantly higher F-scores and better translation performance.

Efficient Path Counting Transducers for Minimum Bayes-Risk Decoding of Statistical Machine Translation Lattices
Graeme Blackwood, Adrià de Gispert and William Byrne
This paper presents an efficient implementation of linearized lattice minimum Bayes-risk decoding using weighted finite state transducers. We introduce transducers to efficiently count lattice paths containing n-grams and use these to gather the required statistics. We show that these procedures can be implemented exactly through simple transformations of word sequences to sequences of n-grams. This yields a novel implementation of lattice minimum Bayes-risk decoding which is fast and exact even for very large lattices.

Short Talks: Discourse and Generation, 11:55–13:15, Venue A, Hall X

"Was It Good? It Was Provocative." Learning the Meaning of Scalar Adjectives
Marie-Catherine de Marneffe, Christopher D. Manning and Christopher Potts
Texts and dialogues often express information indirectly. For instance, speakers' answers to yes/no questions do not always straightforwardly convey a 'yes' or 'no' answer. The intended reply is clear in some cases (Was it good? It was great!) but uncertain in others (Was it acceptable? It was unprecedented.). In this paper, we present methods for interpreting the answers to questions like these which involve scalar modifiers. We show how to ground scalar modifier meaning based on data collected from the Web. We learn scales between modifiers and infer the extent to which a given answer conveys 'yes' or 'no'. To evaluate the methods, we collected examples of question-answer pairs involving scalar modifiers from CNN transcripts and the Dialog Act corpus and use response distributions from Mechanical Turk workers to assess the degree to which each answer conveys 'yes' or 'no'. Our experimental results closely match the Turkers' response data, demonstrating that meanings can be learned from Web data and that such meanings can drive pragmatic inference.

The Same-Head Heuristic for Coreference
Micha Elsner and Eugene Charniak
We investigate coreference relationships between NPs with the same head noun. It is relatively common in unsupervised work to assume that such pairs are coreferent– but this is not always true, especially if realistic mention detection is used. We describe the distribution of non-coreferent same-head pairs in news text, and present an unsupervised generative model which learns not to link some same-head NPs using syntactic features, improving precision.

Authorship Attribution Using Probabilistic Context-Free Grammars
Sindhu Raghavan, Adriana Kovashka and Raymond Mooney
In this paper, we present a novel approach for authorship attribution, the task of identifying the author of a document, using probabilistic context-free grammars. Our approach involves building a probabilistic context-free grammar for each author and using this grammar as a language model

for classification. We evaluate the performance of our method on a wide range of datasets to demonstrate its efficacy.

The Impact of Interpretation Problems on Tutorial Dialogue

Myroslava O. Dzikovska, Johanna D. Moore, Natalie Steinhauser and Gwendolyn Campbell

Supporting natural language input may improve learning in intelligent tutoring systems. However, interpretation errors are unavoidable and require an effective recovery policy. We describe an evaluation on an error recovery policy in the Beetle II tutorial dialogue system and discuss how different types of interpretation problems affect learning gain and user satisfaction. In particular, the problems arising from student use of non-standard terminology appear to have negative consequences. We argue that existing strategies for dealing with terminology problems are insufficient and that improving such strategies is important in future ITS research.

Importance-Driven Turn-Bidding for Spoken Dialogue Systems

Ethan Selfridge and Peter Heeman

Current turn-taking approaches for spoken dialogue systems rely on the speaker releasing the turn before the other can take it. This reliance results in restricted interactions that can lead to inefficient dialogues. In this paper we present a model we refer to as Importance-Driven Turn-Bidding that treats turn-taking as a negotiative process. Each conversant bids for the turn based on the importance of the intended utterance, and Reinforcement Learning is used to indirectly learn this parameter. We find that Importance-Driven Turn-Bidding performs better than two current turn-taking approaches in an artificial collaborative slot-filling domain. The negotiative nature of this model creates efficient dialogues, and supports the improvement of mixed-initiative interaction.

The Prevalence of Descriptive Referring Expressions in News and Narrative

Raquel Hervas and Mark Finlayson

Generating referring expressions is a key step in Natural Language Generation. Researchers have focused almost exclusively on generating distinctive referring expressions, that is, referring expressions that uniquely identify their intended referent. While undoubtedly one of their most important functions, referring expressions can be more than distinctive. In particular, descriptive referring expressions – those that provide additional information not required for distinction – are critical to fluent, efficient, well-written text. We present a corpus analysis in which approximately one-fifth of 7,207 referring expressions in 24,422 words of news and narrative are descriptive. These data show that if we are ever to fully master natural language generation, especially for the genres of news and narrative, researchers will need to devote more attention to understanding how to generate descriptive, and not just distinctive, referring expressions.

Preferences versus Adaptation during Referring Expression Generation

Martijn Goudbeek and Emiel Krahmer

Current Referring Expression Generation algorithms rely on domain dependent preferences for both content selection and linguistic realization. We present two experiments showing that human speakers may opt for dispreferred properties and dispreferred modifier orderings when these were salient in a preceding interaction (without speakers being consciously aware of this). We discuss the impact of these findings for current generation algorithms.

Entity-Based Local Coherence Modelling Using Topological Fields

Jackie Chi Kit Cheung and Gerald Penn

One goal of natural language generation is to produce coherent text that presents information in a logical order. In this paper, we show that topological fields, which model high-level clausal structure, are an important component of local coherence in German. First, we show in a sentence ordering experiment that topological field information improves the entity grid model of Barzilay and Lapata (2008) more than grammatical role and simple clausal order information do, partic-

ularly when manual annotations of this information are not available. Then, we incorporate the model enhanced with topological fields into a natural language generation system that generates constituent orders for German text, and show that the added coherence component improves performance slightly, though not statistically significantly.

Short Talks: Psycholinguistics, Resources, and MT Evaluation, 11:55–13:15, Venue A, Hall IX

Cognitively Plausible Models of Human Language Processing
Frank Keller
We pose the development of cognitively plausible models of human language processing as a challenge for computational linguistics. Existing models can only deal with isolated phenomena (e.g., garden paths) on small, specifically selected data sets. The challenge is to build models that integrate multiple aspects of human language processing at the syntactic, semantic, and discourse level. Like human language processing, these models should be incremental, predictive, broad coverage, and robust to noise. This challenge can only be met if standardized data sets and evaluation measures are developed.

Syntactic and Semantic Factors in Processing Difficulty: An Integrated Measure
Jeff Mitchell, Mirella Lapata, Vera Demberg and Frank Keller
The analysis of reading times can provide insights into the processes that underlie language comprehension, with longer reading times indicating greater cognitive load. There is evidence that the language processor is highly predictive, such that prior context allows upcoming linguistic material to be anticipated. Previous work has investigated the contributions of semantic and syntactic contexts in isolation, essentially treating them as independent factors. In this paper we analyze reading times in terms of a single predictive measure which integrates a model of semantic composition with an incremental parser and a language model.

The Manually Annotated Sub-Corpus: A Community Resource for and by the People
Nancy Ide, Collin Baker, Christiane Fellbaum and Rebecca Passonneau
The Manually Annotated Sub-Corpus (MASC) project provides data and annotations to serve as the base for a community-wide annotation effort of a subset of the American National Corpus. The MASC infrastructure enables the incorporation of contributed annotations into a single, usable format that can then be analyzed as it is or transduced to any of a variety of other formats. MASC includes data from a much wider variety of genres than existing multiply-annotated corpora of English, and the project is committed to a fully open model of distribution, without restriction, for all data and annotations produced or contributed. As such, MASC is the first large-scale, open, community-based effort to create much needed language resources for NLP. This paper describes the MASC project, its corpus and annotations, and serves as a call for contributions of data and annotations from the language processing community.

Correcting Errors in a Treebank Based on Synchronous Tree Substitution Grammar
Yoshihide Kato and Shigeki Matsubara
This paper proposes a method of correcting annotation errors in a treebank. By using a synchronous grammar, the method transforms parse trees containing annotation errors into the ones whose errors are corrected. The synchronous grammar is automatically induced from the treebank. We report an experimental result of applying our method to the Penn Treebank. The result demonstrates that our method corrects syntactic annotation errors with high precision.

Rebanking CCGbank for Improved NP Interpretation
Matthew Honnibal, James R. Curran and Johan Bos
Once released, treebanks tend to remain unchanged despite any shortcomings in their depth of

linguistic analysis or coverage of specific phenomena. Instead, separate resources are created to address such problems. In this paper we show how to improve the quality of a treebank, by integrating resources and implementing improved analyses for specific constructions. We demonstrate this "rebanking" process by creating an updated version of CCGbank that includes the predicate-argument structure of both verbs and nouns, base-NP brackets, verb-particle constructions, and restrictive and non-restrictive nominal modifiers; and evaluate the impact of these changes on a statistical parser.

BabelNet: Building a Very Large Multilingual Semantic Network
Roberto Navigli and Simone Paolo Ponzetto
In this paper we present BabelNet – a very large, wide-coverage multilingual semantic network. The resource is automatically constructed by means of a methodology that integrates lexicographic and encyclopedic knowledge from WordNet and Wikipedia. In addition Machine Translation is also applied to enrich the resource with lexical information for all languages. We conduct experiments on new and existing gold-standard datasets to show the high quality and coverage of the resource.

Evaluating Machine Translations Using mNCD
Marcus Dobrinkat, Tero Tapiovaara, Jaakko Väyrynen and Kimmo Kettunen
This paper introduces mNCD, a method for automatic evaluation of machine translations. The measure is based on normalized compression distance (NCD), a general information theoretic measure of string similarity, and flexible word matching provided by stemming and synonyms. The mNCD measure outperforms NCD in system-level correlation to human judgments in English.

Tackling Sparse Data Issue in Machine Translation Evaluation
Ondřej Bojar, Kamil Kos and David Mareček
We illustrate and explain problems of n-grams-based machine translation (MT) metrics (e.g. BLEU) when applied to morphologically rich languages such as Czech. A novel metric SemPOS based on the deep-syntactic representation of the sentence tackles the issue and retains the performance for translation to English as well.

Short Talks: Semantics 2, 11:55–13:15, Venue B, Lecture Hall 3

Exemplar-Based Models for Word Meaning in Context
Katrin Erk and Sebastian Pado
This paper describes ongoing work on distributional models for word meaning in context. We abandon the usual one-vector-per-word paradigm in favor of an exemplar model that activates only relevant occurrences. On a paraphrasing task, we find that a simple exemplar model outperforms more complex state-of-the-art models.

Fully Unsupervised Core-Adjunct Argument Classification
Omri Abend and Ari Rappoport
The core-adjunct argument distinction is a basic one in the theory of argument structure. The task of distinguishing between the two has strong relations to various basic NLP tasks such as syntactic parsing, semantic role labeling and subcategorization acquisition. This paper presents a novel unsupervised algorithm for the task that uses no supervised models, utilizing instead state-of-the-art syntactic induction algorithms. This is the first work to tackle this task in a fully unsupervised scenario.

A Structured Model for Joint Learning of Argument Roles and Predicate Senses
Yotaro Watanabe, Masayuki Asahara and Yuji Matsumoto
In predicate-argument structure analysis, it is important to capture non-local dependencies among arguments and inter-dependencies between the sense of a predicate and the semantic roles of its

arguments. However, no existing approach explicitly handles both non-local dependencies and semantic dependencies between predicates and arguments. In this paper we propose a structured model that overcomes the limitation of existing approaches; the model captures both types of dependencies simultaneously by introducing four types of factors including a global factor type capturing non-local dependencies among arguments and a pairwise factor type capturing local dependencies between a predicate and an argument. In experiments the proposed model achieved competitive results compared to the state-of-the-art systems without applying any feature selection procedure.

Semantics-Driven Shallow Parsing for Chinese Semantic Role Labeling
Weiwei Sun
One deficiency of current shallow parsing based Semantic Role Labeling (SRL) methods is that syntactic chunks are too small to effectively group words. To partially resolve this problem, we propose semantics-driven shallow parsing, which takes into account both syntactic structures and predicate-argument structures. We also introduce several new "path" features to improve shallow parsing based SRL method. Experiments indicate that our new method obtains a significant improvement over the best reported Chinese SRL result.

Towards Open-Domain Semantic Role Labeling
Danilo Croce, Cristina Giannone, Paolo Annesi and Roberto Basili
Current Semantic Role Labeling technologies are based on inductive algorithms trained over large scale repositories of annotated examples. Frame-based systems currently make use of the FrameNet database but fail to show suitable generalization capabilities in out-of-domain scenarios. In this paper, a state-of-art system for frame-based SRL is extended through the encapsulation of a distributional model of semantic similarity. The resulting argument classification model promotes a simpler feature space that limits the potential overfitting effects. The large scale empirical study here discussed confirms that state-of-art accuracy can be obtained for out-of-domain evaluations.

Collocation Extraction beyond the Independence Assumption
Gerlof Bouma
In this paper we start to explore two-part collocation extraction association measures that do not estimate expected probabilities on the basis of the independence assumption. We propose two new measures based upon the well-known measures of mutual information and pointwise mutual information. Expected probabilities are derived from automatically trained Aggregate Markov Models. On three collocation gold standards, we find the new association measures vary in their effectiveness.

Automatic Collocation Suggestion in Academic Writing
Jian-Cheng Wu, Yu-Chia Chang, Teruko Mitamura and Jason S. Chang
In recent years, collocation has been widely acknowledged as an essential characteristic to distinguish native speakers from non-native speakers. Research on academic writing has also shown that collocations are not only common but serve a particularly important discourse function within the academic community. In our study, we propose a machine learning approach to implementing an online collocation writing assistant. We use a data-driven classifier to provide collocation suggestions to improve word choices, based on the result of classification. The system generates and ranks suggestions to assist learners' collocation usages in their academic writing with satisfactory results.

A Bayesian Method for Robust Estimation of Distributional Similarities
Jun'ichi Kazama, Stijn De Saeger, Kow Kuroda, Masaki Murata and Kentaro Torisawa
Existing word similarity measures are not robust to data sparseness since they rely only on the point estimation of words' context profiles obtained from a limited amount of data. This paper proposes a Bayesian method for robust distributional word similarities. The method uses a distri-

bution of context profiles obtained by Bayesian estimation and takes the expectation of a base similarity measure under that distribution. When the context profiles are multinomial distributions, the priors are Dirichlet, and the base measure is the Bhattacharyya coefficient, we can derive an analytical form that allows efficient calculation. For the task of word similarity estimation using a large amount of Web data in Japanese, we show that the proposed measure gives better accuracies than other well-known similarity measures.

Short Talks: Information Retrieval, Extraction, and Ontologies, 11:55–13:15, Venue B, Lecture Hall 4

Recommendation in Internet Forums and Blogs
Jia Wang, Qing Li, Yuanzhu Peter Chen and Zhangxi Lin
The variety of engaging interactions among users in social medial distinguishes it from traditional Web media. Such a feature should be utilized while attempting to provide intelligent services to social media participants. In this article, we present a framework to recommend relevant information in Internet forums and blogs using user comments, one of the most representative of user behaviors in online discussion. When incorporating user comments, we consider structural, semantic, and authority information carried by them. One of the most important observation from this work is that semantic contents of user comments can play a fairly different role in a different form of social media. When designing a recommendation system for this purpose, such a difference must be considered with caution.

Event-Based Hyperspace Analogue to Language for Query Expansion
Tingxu Yan, Tamsin Maxwell, Dawei Song, Yuexian Hou and Peng Zhang
Bag-of-words approaches to information retrieval (IR) are effective but assume independence between words. The Hyperspace Analogue to Language (HAL) is a cognitively motivated and validated semantic space model that captures statistical dependencies between words by considering their co-occurrences in a surrounding window of text. HAL has been successfully applied to query expansion in IR, but has several limitations, including high processing cost and use of distributional statistics that do not exploit syntax. In this paper, we pursue two methods for incorporating syntactic-semantic information from textual 'events' into HAL. We build the HAL space directly from events to investigate whether processing costs can be reduced through more careful definition of word co-occurrence, and improve the quality of the pseudo-relevance feedback by applying event information as a constraint during HAL construction. Both methods significantly improve performance results in comparison with original HAL, and interpolation of HAL and relevance model expansion outperforms either method alone.

Learning Phrase-Based Spelling Error Models from Clickthrough Data
Xu Sun, Jianfeng Gao, Daniel Micol and Chris Quirk
This paper explores the use of clickthrough data for query spelling correction. First, large amounts of query-correction pairs are derived by analyzing users' query reformulation behavior encoded in the clickthrough data. Then, a phrase-based error model that accounts for the transformation probability between multi-term phrases is trained and integrated into a query speller system. Experiments are carried out on a human-labeled data set. Results show that the system using the phrase-based error model outperforms significantly its baseline systems.

Inducing Domain-Specific Semantic Class Taggers from (Almost) Nothing
Ruihong Huang and Ellen Riloff
This research explores the idea of inducing domain-specific semantic class taggers using only a domain-specific text collection and seed words. The learning process begins by inducing a classifier that only has access to contextual features, forcing it to generalize beyond the seeds. The contextual classifier then labels new instances, to expand and diversify the training set. Next, a

cross-category bootstrapping process simultaneously trains a suite of classifiers for multiple semantic classes. The positive instances for one class are used as negative instances for the others in an iterative bootstrapping cycle. We also explore a one-semantic-class-per-discourse heuristic, and use the classifiers to dynamically create semantic features. We evaluate our approach by inducing six semantic taggers from a collection of veterinary medicine message board posts.

Learning 5000 Relational Extractors
Raphael Hoffmann, Congle Zhang and Daniel S. Weld
Many researchers are trying to use information extraction (IE) to create large-scale knowledge bases from natural language text on the Web. However, the primary approach (supervised learning of relation-specific extractors) requires manually-labeled training data for each relation and doesn't scale to the thousands of relations encoded in Web text. This paper presents WPE, a self-supervised, relation-specific IE system which learns 5025 relations – more than an order of magnitude greater than any previous approach – with an average F1 score of 61%. Crucial to WPE's performance is an automated system for dynamic lexicon learning, which allows it to learn accurately from heuristically-generated training data, which is often noisy and sparse.

Unsupervised Ontology Induction from Text
Hoifung Poon and Pedro Domingos
Extracting knowledge from unstructured text is a long-standing goal of NLP. Although learning approaches to many of its subtasks have been developed (e.g., parsing, taxonomy induction, information extraction), all end-to-end solutions to date require heavy supervision and/or manual engineering, limiting their scope and scalability. We present OntoUSP, a system that induces and populates a probabilistic ontology using only dependency-parsed text as input. OntoUSP builds on the USP unsupervised semantic parser by jointly forming ISA and IS-PART hierarchies of lambda-form clusters. The ISA hierarchy allows more general knowledge to be learned, and the use of smoothing for parameter estimation. We evaluate OntoUSP by using it to extract a knowledge base from biomedical abstracts and answer questions. OntoUSP improves on the recall of USP by 47% and greatly outperforms previous state-of-the-art approaches.

Automatically Generating Term Frequency Induced Taxonomies
Karin Murthy, Tanveer A Faruquie, L Venkata Subramaniam, Hima Prasad K and Mukesh Mohania
We propose a novel method to automatically acquire a term-frequency-based taxonomy from a corpus using an unsupervised method. A term-frequency-based taxonomy is useful for application domains where the frequency with which terms occur on their own and in combination with other terms imposes a natural term hierarchy. We highlight an application for our approach and demonstrate its effectiveness and robustness in extracting knowledge from real-world data.

Complexity Assumptions in Ontology Verbalisation
Richard Power
We describe the strategy currently pursued for verbalising OWL ontologies by sentences in Controlled Natural Language (i.e., combining *generic* rules for realising logical patterns with *ontology-specific* lexicons for realising atomic terms for individuals, classes, and properties) and argue that its success depends on assumptions about the complexity of terms and axioms in the ontology. We then show, through analysis of a corpus of ontologies, that although these assumptions could in principle be violated, they are overwhelmingly respected in practice by ontology developers.

Poster Session 1, 13:15–15:00, Venue A, Foyer

Word Alignment with Synonym Regularization
Hiroyuki Shindo, Akinori Fujino and Masaaki Nagata

We present a novel framework for word alignment that incorporates synonym knowledge collected from monolingual linguistic resources in a bilingual probabilistic model. Synonym information is helpful for word alignment because we can expect a synonym to correspond to the same word in a different language. We design a generative model for word alignment that uses synonym information as a regularization term. The experimental results show that our proposed method significantly improves word alignment quality.

Better Filtration and Augmentation for Hierarchical Phrase-Based Translation Rules
Zhiyang Wang, Yajuan Lv, Qun Liu and Young-Sook Hwang
This paper presents a novel filtration criterion to restrict the rule extraction for the hierarchical phrase-based translation model, where a bilingual but relaxed well-formed dependency restriction is used to filter out bad rules. Furthermore, a new feature which describes the regularity that the source/target dependency edge triggers the target/source word is also proposed. Experimental results show that, the new criteria weeds out about 40% rules while with translation performance improvement, and the new feature brings another improvement to the baseline system, especially on larger corpus.

Fixed Length Word Suffix for Factored Statistical Machine Translation
Narges Sharif Razavian and Stephan Vogel
Factored Statistical Machine Translation extends the Phrase Based SMT model by allowing each word to be a vector of factors. Experiments have shown effectiveness of many factors, including the Part of Speech tags in improving the grammaticality of the output. However, high quality part of speech taggers are not available in open domain for many languages. In this paper we used fixed length word suffix as a new factor in the Factored SMT to replace the part of speech tag factors, and were able to achieve significant improvements in three set of experiments: large NIST Arabic to English system, medium WMT Spanish to English system, and small TRANSTAC English to Iraqi system.

Unsupervised Discourse Segmentation of Documents with Inherently Parallel Structure
Minwoo Jeong and Ivan Titov
Documents often have inherently parallel structure: they may consist of a text and commentaries, or an abstract and a body, or parts presenting alternative views on the same problem. Revealing relations between the parts by jointly segmenting and predicting links between the segments, would help to visualize such documents and construct friendlier user interfaces. To address this problem, we propose an unsupervised Bayesian model for joint discourse segmentation and alignment. We apply our method to the "English as a second language" podcast dataset where each episode is composed of two parallel parts: a story and an explanatory lecture. The predicted topical links uncover hidden relations between the stories and the lectures. In this domain, our method achieves competitive results, rivaling those of a previously proposed supervised technique.

Coreference Resolution with Reconcile
Veselin Stoyanov, Claire Cardie, Nathan Gilbert, Ellen Riloff, David Buttler and David Hysom
Despite the existence of several noun phrase coreference resolution data sets as well as several formal evaluations on the task, it remains frustratingly difficult to compare results across different coreference resolution systems. This is due to the high cost of implementing a complete end-to-end coreference resolution system, which often forces researchers to substitute available gold-standard information in lieu of implementing a module that would compute that information. Unfortunately, this leads to inconsistent and often unrealistic evaluation scenarios. With the aim to facilitate consistent and realistic experimental evaluations in coreference resolution, we present Reconcile, an infrastructure for the development of learning-based noun phrase (NP) coreference resolution systems. Reconcile is designed to facilitate the rapid creation of coreference resolution systems, easy implementation of new feature sets and approaches to coreference resolution, and

empirical evaluation of coreference resolvers across a variety of benchmark data sets and standard scoring metrics. We describe Reconcile and present experimental results showing that Reconcile can be used to create a coreference resolver that achieves performance comparable to state-of-the-art systems on six benchmark data sets.

Predicate Argument Structure Analysis Using Transformation Based Learning
Hirotoshi Taira, Sanae Fujita and Masaaki Nagata
Maintaining high annotation consistency in large corpora is crucial for statistical learning; however, such work is hard, especially for tasks containing semantic elements. This paper describes predicate argument structure analysis using transformation-based learning. An advantage of transformation-based learning is the readability of learned rules. A disadvantage is that the rule extraction procedure is time-consuming. We present incremental-based, transformation-based learning for semantic processing tasks. As an example, we deal with Japanese predicate argument analysis and show some tendencies of annotators for constructing a corpus with our method.

Improving Chinese Semantic Role Labeling with Rich Syntactic Features
Weiwei Sun
Developing features has been shown crucial to advancing the state-of-the-art in Semantic Role Labeling (SRL). To improve Chinese SRL, we propose a set of additional features, some of which are designed to better capture structural information. Our system achieves 93.49 F-measure, a significant improvement over the best reported performance 92.0. We are further concerned with the effect of parsing in Chinese SRL. We empirically analyze the two-fold effect, grouping words into constituents and providing syntactic information. We also give some preliminary linguistic explanations.

Translation 2, 15:00–16:15, Venue A, Aula

Exploring Syntactic Structural Features for Sub-Tree Alignment Using Bilingual Tree Kernels
Jun Sun, Min Zhang and Chew Lim Tan
We propose Bilingual Tree Kernels (BTKs) to capture the structural similarities across a pair of syntactic translational equivalences and apply BTKs to sub-tree alignment along with some plain features. Our study reveals that the structural features embedded in a bilingual parse tree pair are very effective for sub-tree alignment and the bilingual tree kernels can well capture such features. The experimental results show that our approach achieves a significant improvement on both gold standard tree bank and automatically parsed tree pairs against a heuristic similarity based method. We further apply the sub-tree alignment in machine translation with two methods. It is suggested that the sub-tree alignment benefits both phrase and syntax based systems by relaxing the constraint of the word alignment.

Discriminative Pruning for Discriminative ITG Alignment
Shujie Liu, Chi-Ho Li and Ming Zhou
While Inversion Transduction Grammar (ITG) has regained more and more attention in recent years, it still suffers from the major obstacle of speed. We propose a discriminative ITG pruning framework using Minimum Error Rate Training and various features from previous work on ITG alignment. Experiment results show that it is superior to all existing heuristics in ITG pruning. On top of the pruning framework, we also propose a discriminative ITG alignment model using hierarchical phrase pairs, which improves both F-score and Bleu score over the baseline alignment system of GIZA++.

Fine-Grained Tree-to-String Translation Rule Extraction
Xianchao Wu, Takuya Matsuzaki and Jun'ichi Tsujii
Tree-to-string translation rules are widely used in linguistically syntax-based Statistical Machine

Translation systems. In this paper, we propose the usage of deep syntactic information to obtain fine-grained translation rules. A Head-driven Phrase Structure Grammar (HPSG) parser is used to obtain the deep syntactic information of an English sentence, which includes a fine-grained description of the syntactic property and a semantic representation of the sentence. We extract fine-grained rules from aligned HPSG tree/forest-string pairs and apply them to our tree-to-string and string-to-tree systems. Extensive experiments on large-scale bidirectional Japanese-English translations testified the effectiveness of our proposal.

Parsing 2, 15:00–16:15, Venue A, Hall X

Accurate Context-Free Parsing with Combinatory Categorial Grammar

Timothy A. D. Fowler and Gerald Penn

The definition of combinatory categorial grammar (CCG) in the literature varies quite a bit from author to author. However, the differences between the definitions are important in terms of the language classes of each CCG. We prove that a wide range of CCGs are strongly context-free, including the CCG of CCGbank and of the parser of Clark and Curran (2007). In light of these new results, we train the PCFG parser of Petrov and Klein (2007) on CCGbank and achieve state of the art results in supertagging accuracy, PARSEVAL measures and dependency accuracy.

Faster Parsing by Supertagger Adaptation

Jonathan K. Kummerfeld, Jessika Roesner, Tim Dawborn, James Haggerty, James R. Curran and Stephen Clark

We propose a novel self-training method for a parser which uses a lexicalised grammar and supertagger, focusing on increasing the speed of the parser rather than its accuracy. The idea is to train the supertagger on large amounts of parser output, so that the supertagger can learn to supply the supertags that the parser will eventually choose as part of the highest-scoring derivation. Since the supertagger supplies fewer supertags overall, the parsing speed is increased. We demonstrate the effectiveness of the method using a CCG supertagger and parser, obtaining significant speed increases on newspaper text with no loss in accuracy. We also show that the method can be used to adapt the CCG parser to new domains, obtaining accuracy and speed improvements for Wikipedia and biomedical text.

Using Smaller Constituents Rather Than Sentences in Active Learning for Japanese Dependency Parsing

Manabu Sassano and Sadao Kurohashi

We investigate active learning methods for Japanese dependency parsing. We propose active learning methods of using partial dependency relations in a given sentence for parsing and evaluate their effectiveness empirically. Furthermore, we utilize syntactic constraints of Japanese to obtain more labeled examples from precious labeled ones that annotators give. Experimental results show that our proposed methods improve considerably the learning curve of Japanese dependency parsing. In order to achieve an accuracy of over 88.3%, one of our methods requires only 34.4% of labeled examples as compared to passive learning.

Morphology, 14:30–15:45, Venue A, Hall IX

Conditional Random Fields for Word Hyphenation

Nikolaos Trogkanis and Charles Elkan

Finding allowable places in words to insert hyphens is an important practical problem. The algorithm that is used most often nowadays has remained essentially unchanged for 25 years. This method is the TeX hyphenation algorithm of Knuth and Liang. We present here a hyphenation method that is clearly more accurate. The new method is an application of conditional random fields. We create new training sets for English and Dutch from the CELEX European lexical resource, and achieve error rates for English of less than 0.1% for correctly allowed hyphens, and

less than 0.01% for Dutch. Experiments show that both the Knuth/Liang method and a leading current commercial alternative have error rates several times higher for both languages.

Enhanced Word Decomposition by Calibrating the Decision Threshold of Probabilistic Models and Using a Model Ensemble
Sebastian Spiegler and Peter A. Flach
This paper demonstrates that the use of ensemble methods and carefully calibrating the decision threshold can significantly improve the performance of machine learning methods for morphological word decomposition. We employ two algorithms which come from a family of generative probabilistic models. The models consider segment boundaries as hidden variables and include probabilities for letter transitions within segments. The advantage of this model family is that it can learn from small datasets and easily generalises to larger datasets. The first algorithm Promodes, which participated in the Morpho Challenge 2009 (an international competition for unsupervised morphological analysis) employs a lower order model whereas the second algorithm Promodes-H is a novel development of the first using a higher order model. We present the mathematical description for both algorithms, conduct experiments on the morphologically rich language Zulu and compare characteristics of both algorithms based on the experimental results.

Word Representations: A Simple and General Method for Semi-Supervised Learning
Joseph Turian, Lev-Arie Ratinov and Yoshua Bengio
If we take an existing supervised NLP system, a simple and general way to improve accuracy is to use unsupervised word representations as extra word features. We evaluate Brown clusters, Collobert and Weston (2008) embeddings, and HLBL (Mnih Hinton, 2009) embeddings of words on both NER and chunking. We use near state-of-the-art supervised baselines, and find that each of the three word representations improves the accuracy of these baselines. We find further improvements by combining different word representations. You can download our word features, for off-the-shelf use in existing NLP systems, as well as our code, here: http://metaoptimize.com/projects/wordreprs/

Sentiment 1, 15:00–16:15, Venue B, Lecture Hall 3

Identifying Text Polarity Using Random Walks
Ahmed Hassan and Dragomir Radev
Automatically identifying the polarity of words is a very important task in Natural Language Processing. It has applications in text classification, text filtering, analysis of product review, analysis of responses to surveys, and mining online discussions. We propose a method for identifying the polarity of words. We apply a Markov random walk model to a large word relatedness graph, producing a polarity estimate for any given word. A key advantage of the model is its ability to accurately and quickly assign a polarity sign and a magnitude to any word. The method could be used both in a semi-supervised settings where a training set of labeled words is used, and in an unsupervised setting where a handful of seeds is used to define the two polarity classes. The method is experimentally tested using a manually labeled set of positive and negative words. It outperform the state of the art methods in the semi-supervised setting. The results in the unsupervised setting is comparable to the best reported values. However, the proposed method is faster and does not need a large corpus.

Sentiment Learning on Product Reviews via Sentiment Ontology Tree
Wei Wei and Jon Atle Gulla
Existing works on sentiment analysis on product reviews suffer from the following limitations: (1) The knowledge of hierarchical relationships of products attributes is not fully utilized. (2) Reviews or sentences mentioning several attributes associated with complicated sentiments are

not dealt with very well. In this paper, we propose a novel HL-SOT approach to labeling a product's attributes and their associated sentiments in product reviews by a Hierarchical Learning (HL) process with a defined Sentiment Ontology Tree (SOT). The empirical analysis against a human-labeled data set demonstrates promising and reasonable performance of the proposed HL-SOT approach. While this paper is mainly on sentiment analysis on reviews of one product, our proposed HL-SOT approach is easily generalized to labeling a mix of reviews of more than one products.

Employing Personal/Impersonal Views in Supervised and Semi-Supervised Sentiment Classification
Shoushan Li, Chu-Ren Huang, Guodong Zhou and Sophia Yat Mei Lee
In this paper, we adopt two views, personal and impersonal views, and systematically employ them in both supervised and semi-supervised sentiment classification. Here, personal views consist of those sentences which directly express speaker's feeling and preference towards a target object while impersonal views focus on statements towards a target object for evaluation. To obtain them, an unsupervised mining approach is proposed. On this basis, an ensemble method and a co-training algorithm are explored to employ the two views in supervised and semi-supervised sentiment classification respectively. Experimental results across eight domains demonstrate the effectiveness of our proposed approach.

Selectional Preferences, 15:00–16:15, Venue B, Lecture Hall 4

A Latent Dirichlet Allocation Method for Selectional Preferences
Alan Ritter, Mausam and Oren Etzioni
The computation of selectional preferences, the admissible argument values for a relation, is a well-known NLP task with broad applicability. We present LDA-SP, which utilizes LinkLDA (Erosheva et. al. 2004) to model selectional preferences. By simultaneously inferring latent topics and topic distributions over relations, LDA-SP combines the benefits of previous approaches: like traditional class-based approaches, it produces human-interpretable classes describing each relation's preferences, but it is competitive with non-class-based methods in predictive power. We compare LDA-SP to several state-of-the-art methods achieving an 85% increase in recall at 0.9 precision over mutual information (Erk 2007). We also evaluate LDA-SP effectiveness at filtering improper applications of inference rules, where we show substantial improvement over Pantel et. al's system (Pantel 2007).

Latent Variable Models of Selectional Preference
Diarmuid Ó Séaghdha
This paper describes the application of so-called "topic models" to selectional preference induction. Three models related to Latent Dirichlet Allocation, a proven method for modelling document-word co-occurrences, are presented and evaluated on datasets of human plausibility judgements. Compared to previously proposed techniques, these models perform very competitively, especially for infrequent predicate-argument combinations where they exceed the quality of Web-scale predictions while using relatively little data.

Improving the Use of Pseudo-Words for Evaluating Selectional Preferences
Nathanael Chambers and Daniel Jurafsky
This paper improves the use of pseudo-words as an evaluation framework for selectional preferences. While pseudo-words originally evaluated word sense disambiguation, they are now commonly used to evaluate selectional preferences. A selectional preference model ranks a set of possible arguments for a verb by their semantic fit to the verb. Pseudo-words serve as a proxy evaluation for these decisions. The evaluation takes an argument of a verb like drive (e.g. car), pairs it with an alternative word (e.g. car/rock), and asks a model to identify the original. This pa-

per studies two main aspects of pseudo-word creation that affect performance results. (1) Pseudo-word evaluations often evaluate only a subset of the words. We show that selectional preferences should instead be evaluated on the data in its entirety. (2) Different approaches to selecting partner words can produce overly optimistic evaluations. We offer suggestions for normalizing these factors and present a simple baseline that outperforms the state-of-the-art by 13% absolute on a newspaper domain.

Translation 3, 16:45–18:00, Venue A, Aula

Syntax-to-Morphology Mapping in Factored Phrase-Based Statistical Machine Translation from English to Turkish

Reyyan Yeniterzi and Kemal Oflazer

We present a novel scheme to apply factored phrase-based SMT to a language pair with very disparate morphological structures. Our approach relies on syntactic analysis on the source side (English) and then encodes a wide variety of local and non-local syntactic structures as *complex structural tags* which appear as additional factors in the training data. On the target side (Turkish), we only perform morphological analysis and disambiguation but treat the complete complex morphological tag as a factor, instead of separating morphemes. We incrementally explore capturing various syntactic substructures as complex tags on the English side, and evaluate how our translations improve in BLEU scores. Our maximal set of source and target side transformations, coupled with some additional techniques, provide an 39% relative improvement from a baseline 17.08 to 23.78 BLEU, all averaged over 10 training and test sets. Now that the syntactic analysis on the English side is available, we also experiment with more long distance constituent reordering to bring the English constituent order close to Turkish, but find that these transformations do not provide any additional consistent tangible gains when averaged over the 10 sets.

Hindi-to-Urdu Machine Translation through Transliteration

Nadir Durrani, Hassan Sajjad, Alexander Fraser and Helmut Schmid

We present a novel approach to integrate transliteration into Hindi-to-Urdu SMT. We propose two probabilistic models, based on conditional and joint probability formulations, that are novel solutions to the problem. Our models consider both transliteration and translation when translating a particular Hindi word given the context whereas in previous work transliteration is only used for translating OOV (out-of-vocabulary) words. We use transliteration as a tool for disambiguation of Hindi homonyms which can be both translated or transliterated or transliterated differently based on different contexts. We obtain final BLEU scores of 19.35 (conditional probability model) and 19.0 (joint probability model) as compared to 14.30 for a baseline phrase-based system and 16.25 for a system which transliterates OOV words in the baseline system. This indicates that transliteration is useful for more than only translating OOV words for language pairs like Hindi-Urdu.

Training Phrase Translation Models with Leaving-One-Out

Joern Wuebker, Arne Mauser and Hermann Ney

Several attempts have been made to learn phrase translation probabilities for phrase-based statistical machine translation that go beyond pure counting of phrases in word-aligned training data. Most approaches report problems with overfitting. We describe a novel leaving-one-out approach to prevent over-fitting that allows us to train phrase models that show improved translation performance on the WMT08 Europarl German-English task. In contrast to most previous work where phrase models were trained separately from other models used in translation, we include all components such as single word lexica and reordering models in training. Using this consistent training of phrase models we are able to achieve improvements of up to 1.4 points in BLEU. As a side effect, the phrase table size is reduced by more than 80%.

Tagging, 16:45–18:00, Venue A, Hall X

Efficient Staggered Decoding for Sequence Labeling
Nobuhiro Kaji, Yasuhiro Fujiwara, Naoki Yoshinaga and Masaru Kitsuregawa
The Viterbi algorithm is the conventional decoding algorithm most widely adopted for sequence labeling. Viterbi decoding is, however, prohibitively slow when the label set is large, because its time complexity is quadratic in the number of labels. This paper proposes an exact decoding algorithm that overcomes this problem. A novel property of our algorithm is that it efficiently reduces the labels to decoded, while still allowing us to check the optimality of the solution. The experiments on three tasks (POS tagging, joint POS tagging and chunking, and supertagging) show that the new algorithm is several orders of magnitude faster than the basic Viterbi and a state-of-the-art algorithm, CarpeDiem.

Minimized Models and Grammar-Informed Initialization for Supertagging with Highly Ambiguous Lexicons
Sujith Ravi, Jason Baldridge and Kevin Knight
We combine two complementary ideas for learning supertaggers from highly ambiguous lexicons: grammar-informed tag transitions and models minimized via integer programming. Each strategy on its own greatly improves performance over basic expectation-maximization training with a bitag Hidden Markov Model, which we show on the CCGbank and CCG-TUT corpora. The strategies provide further error reductions when combined. We describe a new two-stage integer programming strategy that efficiently deals with the high degree of ambiguity on these datasets while obtaining the full effect of model minimization.

Practical Very Large Scale CRFs
Thomas Lavergne, Olivier Cappé and François Yvon
Conditional Random Fields (CRFs) constitute a popular approach for supervised sequence labelling, notably due to their ability to handle large description spaces and to integrate structural dependency between labels. Taking structure into account typically implies a number of parameters and a computational effort that grows quadratically with the cardinality of the label set. In this paper, we address the issue of training very large CRFs, containing up to hundreds output labels and several billions of features. Efficiency stems here from the sparsity induced by the use of a l1 penalty term. Based on our own implementation, we compare three recent proposal for implementing this regularization strategy. Our experiments demonstrate that very large CRFs can be trained efficiently and that larger models are able to improve the accuracy, while delivering compact parameter sets.

Grammar Formalisms, 16:45–18:00, Venue A, Hall IX

On the Computational Complexity of Dominance Links in Grammatical Formalisms
Sylvain Schmitz
Dominance links were introduced in grammars to model long distance scrambling phenomena, motivating the definition of multiset-valued linear indexed grammars (MLIGs) by Rambow (1994), and inspiring quite a few recent formalisms. It turns out that MLIGs have since been rediscovered and reused in a variety of contexts, and that the complexity of their emptiness problem has become the key to several open questions in computer science. We survey complexity results and open issues on MLIGs and related formalisms, and provide new complexity bounds for some linguistically motivated restrictions.

Optimal Rank Reduction for Linear Context-Free Rewriting Systems with Fan-Out Two
Benoît Sagot and Giorgio Satta
Linear Context-Free Rewriting Systems (LCFRSs) are a grammar formalism capable of modeling

discontinuous phrases. Many parsing applications use LCFRSs where the fan-out (a measure of the discontinuity of phrases) does not exceed 2. We present an efficient algorithm for optimal reduction of the length of production right-hand side in LCFRSs with fan-out at most 2. This results in asymptotical running time improvement for known parsing algorithms for this class.

The Importance of Rule Restrictions in CCG
Marco Kuhlmann, Alexander Koller and Giorgio Satta
Combinatory Categorial Grammar (CCG) is generally construed as a fully lexicalized formalism, where all grammars use one and the same universal set of rules, and cross-linguistic variation is isolated in the lexicon. In this paper, we show that the weak generative capacity of this "pure" form of CCG is strictly smaller than that of CCG with grammar-specific rules, and of other mildly context-sensitive grammar formalisms, including Tree Adjoining Grammar (TAG). Our result also carries over to a multi-modal extension of CCG.

Summarization 1, 16:45–18:00, Venue B, Lecture Hall 3

Automatic Evaluation of Linguistic Quality in Multi-Document Summarization
Emily Pitler, Annie Louis and Ani Nenkova
To date, few attempts have been made to develop and validate methods for automatic evaluation of linguistic quality in text summarization. We present the first systematic assessment of several diverse classes of metrics designed to capture various aspects of well-written text. We train and test linguistic quality models on consecutive years of NIST evaluation data in order to show the generality of results. For grammaticality, the best results come from a set of syntactic features. Focus, coherence and referential clarity are best evaluated by a class of features measuring local coherence on the basis of cosine similarity between sentences, coreference information, and summarization specific features. Our best results are 90% accuracy for pairwise comparisons of competing systems over a test set of several inputs and 70% for ranking summaries of a specific input.

Identifying Non-Explicit Citing Sentences for Citation-Based Summarization.
Vahed Qazvinian and Dragomir Radev
Identifying background (context) information in scientific articles can help scholars understand major contributions in their research area more easily. In this paper, we propose a general framework based on probabilistic inference to extract such context information from scientific papers. We model the sentences in an article and their lexical similarities as a Markov Random Field tuned to detect the patterns that context data create, and employ a Belief Propagation mechanism to detect likely context sentences. We also address the problem of generating surveys of scientific papers. Our experiments show greater pyramid scores for surveys generated using such context information rather than citation sentences alone.

Automatic Generation of Story Highlights
Kristian Woodsend and Mirella Lapata
In this paper we present a joint content selection and compression model for single-document summarization. The model operates over a phrase-based representation of the source document which we obtain by merging information from PCFG parse trees and dependency graphs. Using an integer linear programming formulation, the model learns to select and combine phrases subject to length, coverage and grammar constraints. We evaluate the approach on the task of generating "story highlights"—a small number of brief, self-contained sentences that allow readers to quickly gather information on news stories. Experimental results show that the model's output is comparable to human-written highlights in terms of both grammaticality and content.

Sentiment 2, 16:45–18:00, Venue B, Lecture Hall 4

Sentence and Expression Level Annotation of Opinions in User-Generated Discourse
Cigdem Toprak, Niklas Jakob and Iryna Gurevych
In this paper, we introduce a corpus of consumer reviews from the rateitall and the eopinionswebsites annotated with opinion-related information. We present a two-level annotation scheme. In the first stage, the reviews are analyzed at the sentence level for (i) relevancy to a given topic, and (ii) expressing an evaluation about the topic. In the second stage, on-topic sentences containing evaluations about the topic are further investigated at the expression level for pinpointing the properties (semantic orientation, intensity), and the functional components of the evaluations (opinion terms, targets and holders). We discuss the annotation scheme, the inter-annotator agreement for different subtasks and our observations.

Generating Focused Topic-Specific Sentiment Lexicons
Valentin Jijkoun, Maarten de Rijke and Wouter Weerkamp
We present a method for automatically generating focused and accurate topic-specific subjectivity lexicons that allow users to pin-point subjective on-topic information in a set of relevant documents. We motivate the need for such lexicons in the field of media analysis, describe a bootstrapping method for generating a topic-specific lexicon from a general purpose polarity lexicon, and evaluate the quality of the generated lexicons both manually and using a TREC Blog track test set for opinionated blog post retrieval. Although the generated lexicons can be an order of magnitude more selective, they maintain, or even improve, the performance of an opinion retrieval system.

Evaluating Multilanguage-Comparability of Subjectivity Analysis Systems
Jungi Kim, Jin-Ji Li and Jong-Hyeok Lee
Subjectivity analysis is a rapidly growing field of study. Along with its applications to various NLP tasks, much work have put efforts into multilingual subjectivity learning from existing resources. Multilingual subjectivity analysis requires language-independent criteria for comparable outcomes across languages. This paper proposes to measure the multilanguage-comparability of subjectivity analysis tools, and provides meaningful comparisons of multilingual subjectivity analysis from various points of view.

ACL 2010 Main Conference Abstracts: Tuesday, July 13

Translation 4, 10:30–11:45, Venue A, Aula

Error Detection for Statistical Machine Translation Using Linguistic Features
Deyi Xiong, Min Zhang and Haizhou Li
Automatic error detection is desired in the post-processing to improve machine translation quality. The previous work is largely based on confidence estimation using system-based features, such as word posterior probabilities calculated from N-best lists or word lattices. We propose to incorporate two groups of linguistic features, which convey information from outside machine translation systems, into error detection: lexical and syntactic features. We use a maximum entropy classifier to predict translation errors by integrating word posterior probability feature and linguistic features. The experimental results show that 1) linguistic features alone outperform word posterior probability based confidence estimation in error detection; and 2) linguistic features can further provide complementary information when combined with word confidence scores, which collectively reduce the classification error rate by 18.52% and improve the F measure by 16.37%.

TrustRank: Inducing Trust in Automatic Translations via Ranking
Radu Soricut and Abdessamad Echihabi
The adoption of Machine Translation technology for commercial applications is hampered by the lack of trust associated with machine-translated output. In this paper, we describe TrustRank, an MT system enhanced with a capability to rank the quality of translation outputs from good to bad. This enables the user to set a quality threshold, granting the user control over the quality of the translations. We quantify the gains we obtain in translation quality, and show that our solution works on a wide variety of domains and language pairs.

Bridging SMT and TM with Translation Recommendation
Yifan He, Yanjun Ma, Josef van Genabith and Andy Way
We propose a translation recommendation framework to integrate Statistical Machine Translation (SMT) output with Translation Memory TM) systems. The framework recommends SMT outputs to a TM user when it predicts that SMT outputs are more suitable for post-editing than the hits provided by the TM. We describe an implementation of this framework using an SVM binary classifier. We exploit methods to fine-tune the classifier and investigate a variety of features of different types. We rely on automatic MT evaluation metrics to approximate human judgements in our experiments. Experimental results show that our system can achieve 0.85 precision at 0.89 recall, excluding exact matches. Furthermore, it is possible for the end-user to achieve a desired balance between precision and recall by adjusting confidence levels.

Information Extraction 2, 10:30–11:45, Venue A, Hall X

On Jointly Recognizing and Aligning Bilingual Named Entities
Yufeng Chen, Chengqing Zong and Keh-Yih Su
We observe that (1) how a given named entity (NE) is translated (i.e., either semantically or phonetically) depends greatly on its associated entity type, and (2) entities within an aligned pair should share the same type. Also, (3) those initially detected NEs are anchors, whose information should be used to give certainty scores when selecting candidates. From this basis, an integrated model is thus proposed in this paper to jointly identify and align bilingual named entities between Chinese and English. It adopts a new mapping type ratio feature (which is the proportion of NE internal tokens that are semantically translated), enforces an entity type consistency constraint, and utilizes additional monolingual candidate certainty factors (based on those NE anchors). The experiments show that this novel approach has substantially raised the type-sensitive F-score of

identified NE-pairs from 68.4% to 81.7% (42.1% F-score imperfection reduction) in our Chinese-English NE alignment task.

Generating Templates of Entity Summaries with an Entity-Aspect Model and Pattern Mining
Peng Li, Jing Jiang and Yinglin Wang
In this paper, we propose a novel approach to automatic generation of summary templates from given collections of summary articles. This kind of summary templates can be useful in various applications. We first develop an entity-aspect LDA model to simultaneously cluster both sentences and words into aspects. We then apply frequent subtree pattern mining on the dependency parse trees of the clustered and labeled sentences to discover sentence patterns that well represent the aspects. Key features of our method include automatic grouping of semantically related sentence patterns and automatic identification of template slots that need to be filled in. We apply our method on five Wikipedia entity categories and compare our method with two baseline methods. Both quantitative evaluation based on human judgment and qualitative comparison demonstrate the effectiveness and advantages of our method.

Comparable Entity Mining from Comparative Questions
Shasha Li, Chin-Yew Lin, Young-In Song and Zhoujun Li
Comparing one thing with another is a typical part of human decision making process. However, it is not always easy to know what to compare and what are the alternatives. To address this difficulty, we present a novel way to automatically mine comparable entities from comparative questions that users posted online. To ensure high precision and high recall, we develop a weakly-supervised bootstrapping method for comparative question identification and comparable entity extraction by leveraging a large online question archive. The experimental results show our method achieves F1-measure of 82.5% in comparative question identification and 83.3% in comparable entity extraction. Both significantly outperform an existing state-of-the-art method.

Student Research Workshop, 10:30–11:45, Venue A, Hall IX

Non-Cooperation in Dialogue
Brian Plüss
This paper presents ongoing research on computational models for non-cooperative dialogue. We start by analysing different levels of cooperation in conversation. Then, inspired by findings from an empirical study, we propose a technique for measuring non-cooperation in political interviews. Finally, we describe a research programme towards obtaining a suitable model and discuss previous accounts for conflictive dialogue, identifying the differences with our work.

Towards Relational POMDPs for Adaptive Dialogue Management
Pierre Lison
Open-ended spoken interactions are typically characterised by both structural complexity and high levels of uncertainty, making dialogue management in such settings a particularly challenging problem. Traditional approaches have focused on providing theoretical accounts for either the uncertainty or the complexity of spoken dialogue, but rarely considered the two issues simultaneously. This paper describes ongoing work on a new approach to dialogue management which attempts to fill this gap. We represent the interaction as a Partially Observable Markov Decision Process (POMDP) over a rich state space incorporating both dialogue, user, and environment models. The tractability of the resulting POMDP can be preserved using a mechanism for dynamically constraining the action space based on prior knowledge over locally relevant dialogue structures. These constraints are encoded in a small set of general rules expressed as a Markov Logic network. The first-order expressivity of Markov Logic enables us to leverage the rich relational structure of the problem and efficiently abstract over large regions of the state and action spaces.

WSD as a Distributed Constraint Optimization Problem
Siva Reddy and Abhilash Inumella
This work models Word Sense Disambiguation (WSD) problem as a Distributed Constraint Optimization Problem (DCOP). To model WSD as a DCOP, we view information from various knowledge sources as constraints. DCOP algorithms have the remarkable property to jointly maximize over a wide range of utility functions associated with these constraints. We show how utility functions can be designed for various knowledge sources. For the purpose of evaluation, we modelled all words WSD as a simple DCOP problem. The results are competitive with state-of-art knowledge based systems.

A Probabilistic Generative Model for an Intermediate Constituency-Dependency Representation
Federico Sangati
We present a probabilistic model extension to the Tesnière Dependency Structure (TDS) framework formulated in (Sangati and Mazza, 2009). This representation incorporates aspects from both constituency and dependency theory. In addition, it makes use of junction structures to handle coordination constructions. We test our model on parsing the English Penn WSJ treebank using a re-ranking framework. This technique allows us to efficiently test our model without needing a specialized parser, and to use the standard evaluation metric on the original Phrase Structure version of the treebank. We obtain encouraging results: we achieve a small improvement over state-of-the-art results when re-ranking a small number of candidate structures, on all the evaluation metrics except for chunking.

Sentiment Translation through Lexicon Induction
Christian Scheible
The translation of sentiment information is a task from which sentiment analysis systems can benefit. We present a novel, graph-based approach using SimRank, a well-established vertex similarity algorithm to transfer sentiment information between a source language and a target language graph. We evaluate this method in comparison with SO-PMI (Turney, 2002) against a test set which was annotated by 9 human judges. To compare the two methods to the human raters, we first examine the correlation coefficients. The correlation coefficients between the automatic systems and the human ratings, the two methods yield correlation coefficients which are not significantly different, thus SO and SR have about the same performance on this broad measure. Since many adjectives do not express sentiment at all, the correct categorization of neutral adjectives is just as important as the scalar rating. Thus, we divide the adjectives into three categories – positive, neutral, and negative – with a varying threshold between those categories. Overall, SimRank performs better than SO-PMI for a plausible neutral threshold on the human ratings.

Unsupervised Search for The Optimal Segmentation for Statistical Machine Translation
Coşkun Mermer and Ahmet Afşın Akın
We tackle the previously unaddressed problem of unsupervised determination of the optimal morphological segmentation for statistical machine translation (SMT) and propose a segmentation metric that takes into account both sides of the SMT training corpus. We formulate the objective function as the posterior probability of the training corpus according to a generative segmentation-translation model. We describe how the IBM Model-1 translation likelihood can be computed incrementally between adjacent segmentation states for efficient computation. Submerging the proposed segmentation method in an SMT task from morphologically-rich Turkish to English does not exhibit the expected improvement in translation BLEU scores and confirms the robustness of phrase-based SMT to translation unit combinatorics. A positive outcome of this work is the described modification to the sequential search algorithm of Morfessor (Creutz and Lagus, 2007) that enables arbitrary-fold parallelization of the computation, which unexpectedly improves the translation performance as measured by BLEU.

How Spoken Language Corpora can Refine Current Speech Motor Training Methodologies
Daniil Umanski and Federico Sangati
The growing availability of spoken language corpora presents new opportunities for enriching the methodologies of speech and language therapy. In this paper, we present a novel approach for constructing speech motor exercises, based on linguistic knowledge extracted from spoken language corpora. In our study with the Dutch Spoken Corpus, syllabic inventories were obtained by means of automatic syllabification of the spoken language data. Our experimental syllabification method exhibited a reliable performance, and allowed for the acquisition of syllabic tokens from the corpus. Consequently, the syllabic tokens were integrated in a tool for clinicians, a result which holds the potential of contributing to the current state of speech motor training methodologies.

Resources, 10:30–11:45, Venue B, Lecture Hall 3

Towards Robust Multi-Tool Tagging. An OWL/DL-Based Approach
Christian Chiarcos
This paper describes a series of experiments to test the hypothesis that the parallel application of multiple NLP tools and the integration of their results improves the correctness and robustness of the resulting analysis. We describe how annotations created by seven NLP tools are mapped onto tool-independent descriptions by means of an ontology of linguistic annotations, and how a majority vote and ontological consistency constraints can be used to integrate multiple alternative analyses of the same token in a consistent way. For morphosyntactic (parts of speech) and morphological annotations of three German corpora, the resulting set of automatically determined ontological descriptions is evaluated in comparison to the (ontological representation of the) existing reference annotation.

Temporal Information Processing of a New Language: Fast Porting with Minimal Resources
Francisco Costa and António Branco
We describe the semi-automatic adaptation of a TimeML annotated corpus from English to Portuguese, a language for which TimeML annotated data was not available yet. In order to validate this adaptation, we use the obtained data to replicate some results in the literature that used the original English data. The fact that comparable results are obtained indicates that our approach can be used successfully to rapidly create semantically annotated resources for new languages.

A Taxonomy, Dataset, and Classifier for Automatic Noun Compound Interpretation
Stephen Tratz and Eduard Hovy
The automatic interpretation of noun-noun compounds is an important subproblem within many natural language processing applications and is an area of increasing interest. The problem is difficult, with disagreement regarding the number and nature of the relations, low inter-annotator agreement, and limited annotated data. In this paper, we present a novel taxonomy of relations that integrates previous relations, the largest publicly-available annotated dataset, and a supervised classification method for automatic noun compound interpretation.

Discourse 1, 10:30–11:45, Venue B, Lecture Hall 4

Models of Metaphor in NLP
Ekaterina Shutova
Automatic processing of metaphor can be clearly divided into two subtasks: metaphor recognition (distinguishing between literal and metaphorical language in a text) and metaphor interpretation (identifying the intended literal meaning of a metaphorical expression). Both of them have been repeatedly addressed in NLP. This paper is the first comprehensive and systematic review of the existing computational models of metaphor, the issues of metaphor annotation in corpora and the available resources.

A Game-Theoretic Model of Metaphorical Bargaining
Beata Beigman Klebanov and Eyal Beigman
We present a game-theoretic model of bargaining over a metaphor in the context of political communication, find its equilibrium, and use it to rationalize observed linguistic behavior. We argue that game theory is well suited for modeling discourse as a dynamic resulting from a number of conflicting pressures, and suggest applications of interest to computational linguists.

Kernel Based Discourse Relation Recognition with Temporal Ordering Information
WenTing Wang, Jian Su and Chew Lim Tan
Syntactic knowledge is important for dis-course relation recognition. Yet only heuristically selected flat paths and 2-level production rules have been used to incorporate such information so far. In this paper we propose using tree kernel based approach to automatically mine the syntactic information from the parse trees for discourse analysis, applying kernel function to the tree structures directly. These structural syntactic features, together with other normal flat features are incorporated into our composite kernel to capture diverse knowledge for simultaneous discourse identification and classification for both explicit and implicit relations. The experiment shows tree kernel approach is able to give statistical significant improvements over flat syntactic path feature. We also illustrate that tree kernel approach covers more structure information than the production rules, which allows tree kernel to further incorporate in-formation from a higher dimension space for possible better discrimination. Besides, we further propose to leverage on temporal ordering information to constrain the interpretation of discourse relation, which also demonstrate statistical significant improvements for discourse relation recognition on PDTB 2.0 for both explicit and implicit as well.

Short Talks: Translation and Parsing, 11:55–13:15, Venue A, Aula

Balancing User Effort and Translation Error in Interactive Machine Translation via Confidence Measures
Jesús González Rubio, Daniel Ortiz Martínez and Francisco Casacuberta
This work deals with the application of confidence measures within an interactive-predictive machine translation system in order to reduce human effort. If a small loss in translation quality can be tolerated for the shake of efficiency, user effort can be saved by interactively translating only those initial translations which the confidence measure classifies as incorrect. We apply confidence estimation as a way to achieve a balance between user effort savings and final translation error. Empirical results show that our proposal allows to obtain almost perfect translations while significantly reducing user effort.

Improving Arabic-to-English Statistical Machine Translation by Reordering Post-Verbal Subjects for Alignment
Marine Carpuat, Yuval Marton and Nizar Habash
We study the challenges raised by Arabic verb and subject detection and reordering in Statistical Machine Translation (SMT). We show that post-verbal subject (VS) constructions are hard to translate because they have highly ambiguous reordering patterns when translated to English. In addition, implementing reordering is difficult because the boundaries of VS constructions are hard to detect accurately, even with a state-of-the-art Arabic dependency parser. We therefore propose to reorder VS constructions into SV order for SMT word alignment only. This strategy significantly improves BLEU and TER scores, even on a strong large-scale baseline and despite noisy parses.

Learning Common Grammar from Multilingual Corpus
Tomoharu Iwata, Daichi Mochihashi and Hiroshi Sawada
We propose a corpus-based probabilistic framework to extract hidden common syntax across

languages from non-parallel multilingual corpora in an unsupervised fashion. For this purpose, we assume a generative model for multilingual corpora, where each sentence is generated from a language dependent probabilistic context-free grammar (PCFG), and these PCFGs are generated from a prior grammar that is common across languages. We also develop a variational method for efficient inference. Experiments on a non-parallel multilingual corpus of eleven languages demonstrate the feasibility of the proposed method.

Hierarchical Joint Learning: Improving Joint Parsing and Named Entity Recognition with Non-Jointly Labeled Data

Jenny Rose Finkel and Christopher D. Manning

One of the main obstacles to producing high quality joint models is the lack of jointly annotated data. Joint modeling of multiple natural language processing tasks outperforms single-task models learned from the same data, but still underperforms compared to single-task models learned on the more abundant quantities of available single-task annotated data. In this paper we present a novel model which makes use of additional single-task annotated data to improve the performance of a joint model. Our model utilizes a hierarchical prior to link the feature weights for shared features in several single-task models and the joint model. Experiments on joint parsing and named entity recognition, using the OntoNotes corpus, show that our hierarchical joint model can produce substantial gains over a joint model trained on only the jointly annotated data.

Detecting Errors in Automatically-Parsed Dependency Relations

Markus Dickinson

We outline different methods to detect errors in automatically-parsed dependency corpora, by comparing so-called dependency rules to their representation in the training data and flagging anomalous ones. By comparing each new rule to every relevant rule from training, we can identify parts of parse trees which are likely erroneous. Even the relatively simple methods of comparison we propose show promise for speeding up the annotation process.

Tree-Based Deterministic Dependency Parsing — An Application to Nivre's Method —

Kotaro Kitagawa and Kumiko Tanaka-Ishii

This article describes new model of statistical dependency parsing based on hierarchical tree structurization. Nivre's deterministic model attempts to determine the global sentence structure from a sequence of parsing actions, each of which concerns only two words and their locally relational words, but more of the global structure should be taken into account to decide parsing actions. We solves this problem by applying parsing actions based on tree-based model.All the words necessary for judgment are considered by including words in the trees; the model then chooses the most probable head candidate from each tree. In an evaluation experiment using the Penn Treebank(WSJ), the proposed model achieved higher accuracy than did previous deterministic models. In terms of the ratio of sentences parsed completely, it slightly outperformed McDonald's optimizing method, which takes account of sibling nodes. Although the proposed model's time complexity is O(n2), the experimental results demonstrated an average parsing time not much slower than O(n).

Sparsity in Dependency Grammar Induction

Jennifer Gillenwater, Kuzman Ganchev, João Graça, Fernando Pereira and Ben Taskar

A strong inductive bias is essential in unsupervised grammar induction. We explore a particular sparsity bias in dependency grammars that encourages a small number of unique dependency types. Specifically, we investigate sparsity-inducing penalties on the posterior distributions of parent-child POS tag pairs in the posterior regularization (PR) framework of Graça et al. (2007). In experiments with 12 languages, we achieve substantial gains over the standard expectation maximization (EM) baseline, with average improvement in attachment accuracy of 6.3%. Further, our method outperforms models based on a standard Bayesian sparsity-inducing prior by an av-

erage of 4.9%. On English in particular, we show that our approach improves on several other state-of-the-art techniques.

Top-Down K-Best A* Parsing
Adam Pauls, Dan Klein and Chris Quirk
We propose a top-down algorithm for extracting k-best lists from a parser. Our algorithm, TKA* is a variant of the k-best A* (KA*) algorithm. In contrast to KA*, which performs an inside and outside pass before performing k-best extraction bottom up, TKA* performs only the inside pass before extracting k-best lists top down. TKA* maintains the same optimality and efficiency guarantees of KA*, but is simpler to both specify and implement.

Short Talks: Machine Learning and Statistical Methods, 11:55–13:15, Venue A, Hall X

Simple Semi-Supervised Training of Part-Of-Speech Taggers
Anders Søgaard
Most attempts to train part-of-speech taggers on a mixture of labeled and unlabeled data have failed (Abney, 2008). In this work stacking (Wolpert, 1992) is used to reduce tagging to a classification task. This simplifies semi-supervised training considerably. Our prefered semi-supervised method combines tri-training (Li Zhou, 2005) and disagreement-based co-training. On the Wall Street Journal, we obtain an error reduction of 4.2% with SVMTool (Gimenez Marquez, 2004).

Efficient Optimization of an MDL-Inspired Objective Function for Unsupervised Part-Of-Speech Tagging
Ashish Vaswani, Adam Pauls and David Chiang
The Minimum Description Length (MDL) principle is a method for model selection that trades off between the explanation of the data by the model and the complexity of the model itself. Inspired by the MDL principle, we develop an objective function for generative models that captures the description of the data by the model (log-likelihood) and the description of the model (model size). We also develop a efficient general search algorithm based on the MAP-EM framework to optimize this function. Since recent work has shown that minimizing the model size in a Hidden Markov Model for part-of-speech (POS) tagging leads to higher accuracies, we test our approach by applying it to this problem. The search algorithm involves a simple change to EM and achieves high POS tagging accuracies on both English and Italian data sets

SVD and Clustering for Unsupervised POS Tagging
Michael Lamar, Yariv Maron, Mark Johnson and Elie Bienenstock
We revisit the algorithm of Schütze (1995) for unsupervised part-of-speech tagging. The algorithm uses reduced-rank singular value decomposition followed by clustering to extract latent features from context distributions. As implemented here, it achieves state-of-the-art tagging accuracy at considerably less cost than more recent methods. It can also produce a range of finer-grained taggings, with potential applications to various tasks.

Intelligent Selection of Language Model Training Data
Robert C. Moore and William Lewis
We address the problem of selecting non-domain-specific language model training data to build auxiliary language models for use in tasks such as machine translation. Our approach is based on comparing the cross-entropy, according to domain-specific and non-domain-specifc language models, for each sentence of the text source used to produce the latter language model. We show that this produces better language models, trained on less data, than both random data selection and two other previously proposed methods.

Blocked Inference in Bayesian Tree Substitution Grammars
Trevor Cohn and Phil Blunsom
Learning a tree substitution grammar is very challenging due to derivational ambiguity. Our recent approach used a Bayesian non-parametric model to induce good derivations from treebanked input (Cohn et al., 2009), biasing towards small grammars composed of small generalisable productions. In this paper we present a novel training method for the model using a blocked Metropolis-Hastings sampler in place of the previous method's local Gibbs sampler. The blocked sampler makes considerably larger moves than the local sampler and consequently converges in less time. A core component of the algorithm is a grammar transformation which represents an infinite tree substitution grammar in a finite context free grammar. This enables efficient blocked inference for training and also improves the parsing algorithm. Both algorithms are shown to improve parsing accuracy.

Boosting-Based System Combination for Machine Translation
Tong Xiao, Jingbo Zhu, Muhua Zhu and Huizhen Wang
In this paper, we present a simple and effective method to address the issue of how to generate diversified translation systems from a single Statistical Machine Translation (SMT) engine for system combination. Our method is based on the framework of boosting. First, a sequence of weak translation systems is generated from a baseline system in an iterative manner. Then, a strong translation system is built from the ensemble of these weak translation systems. To adapt boosting to SMT system combination, several key components of the original boosting algorithms are redesigned in this work. We evaluate our method on Chinese-to-English Machine Translation (MT) tasks in three baseline systems, including a phrase-based system, a hierarchical phrase-based system and a syntax-based system. The experimental results on three NIST evaluation test sets show that our method leads to significant improvements in translation accuracy over the baseline systems.

Fine-Grained Genre Classification Using Structural Learning Algorithms
Zhili Wu, Katja Markert and Serge Sharoff
Prior use of machine learning in genre classification used a list of labels as classification categories. However, genre classes are often organised into hierarchies, e.g. covering the subgenres of fiction. In this paper we present a method of using the hierarchy of labels to improve the classification accuracy. As a testbed for this approach we use the Brown Corpus as well as a range of other corpora, including the BNC, HGC and Syracuse. The results are not encouraging: apart from the Brown corpus, the improvements of our structural classifier over the flat one are not statistically significant. We discuss the relation between structural learning performance and the visual and distributional balance of the label hierarchy, suggesting that only balanced hierarchies might profit from structural learning.

Online Generation of Locality Sensitive Hash Signatures
Benjamin Van Durme and Ashwin Lall
Motivated by the recent interest in streaming algorithms for processing large text collections, we revisit the work of Ravichandran et al. (2005) on using the Locality Sensitive Hash (LSH) method of Charikar (2002) to enable fast, approximate comparisons of vector cosine similarity. For the common case of feature updates being additive over a data stream, we show that LSH signatures can be maintained online, without additional approximation error, and with lower memory requirements than when using the standard offline technique.

Short Talks: Question Answering, Entailment and Sentiment, 11:55–13:15, Venue A, Hall IX

Metadata-Aware Measures for Answer Summarization in Community Question Answering
Mattia Tomasoni and Minlie Huang
Our paper presents a framework for automatically processing information coming from community Question Answering (cQA) portals with the purpose of generating a trustful, complete, relevant and succinct summary of answers posted by users. We exploit the metadata intrinsically present in User Generated Content (UGC) to bias automatic multi-document summarization techniques toward high quality information. We adopt a representation of concepts alternative to n-grams and propose two concept-scoring functions based on semantic overlap. Experimental results on data drawn from Yahoo! Answers demonstrate the effectiveness of our method in terms of ROUGE scores. We show that the information contained in the best answers voted by users of cQA portals can be successfully complemented by our method.

Optimizing Question Answering Accuracy by Maximizing Log-Likelihood
Matthias H. Heie, Edward W. D. Whittaker and Sadaoki Furui
In this paper we demonstrate that there is a strong correlation between the Question Answering (QA) accuracy and the log-likelihood of the answer typing component of our statistical QA model. We exploit this observation in a clustering algorithm which optimizes QA accuracy by maximizing the log-likelihood of a set of question-and-answer pairs. Experimental results show that we achieve better QA accuracy using the resulting clusters than by using manually derived clusters.

Generating Entailment Rules from FrameNet
Roni Ben Aharon, Idan Szpektor and Ido Dagan
Many NLP tasks need accurate knowledge for semantic inference. To this end, mostly WordNet is utilized. Yet WordNet is limited, especially for inference between predicates. To help filling this gap, we present an algorithm that generates inference rules between predicates from FrameNet. Our experiment shows that the novel resource is effective and complements WordNet in terms of rule coverage.

Don't 'Have a Clue'? Unsupervised Co-Learning of Downward-Entailing Operators.
Cristian Danescu-Niculescu-Mizil and Lillian Lee
Researchers in textual entailment have begun to consider inferences involving downward-entailing operators, an interesting and important class of lexical items that change the way inferences are made. Recent work proposed a method for learning English downward-entailing operators that requires access to a high-quality collection of negative polarity items (NPIs). However, English is one of the very few languages for which such a list exists. We propose the first approach that can be applied to the many languages for which there is no pre-existing hight-precision database of NPIs. As a case study, we apply our method to Romanian and show that our method yields good results. Also, we perform a cross-linguistic analysis that suggests interesting connections to some findings in linguistic typology.

Vocabulary Choice as an Indicator of Perspective
Beata Beigman Klebanov, Eyal Beigman and Daniel Diermeier
We establish the following characteristics of the task of perspective classification: (a) using term frequencies in a document does not improve classification achieved with absence/presence features; (b) for datasets allowing the relevant comparisons, a small number of top features is found to be as effective as the full feature set and indispensable for the best achieved performance, testifying to the existence of perspective-specific keywords. We relate our findings to research on word frequency distributions and to discourse analytic studies of perspective.

Cross Lingual Adaptation: An Experiment on Sentiment Classifications
Bin Wei and Christopher Pal
In this paper, we study the problem of using an annotated corpus in English for the same natu-

ral language processing task in another language. While various machine translation systems are available, automated translation is still far from perfect. To minimize the noise introduced by translations, we propose to use only key 'reliable" parts from the translations and apply structural correspondence learning (SCL) to find a low dimensional representation shared by the two languages. We perform experiments on an English-Chinese sentiment classification task and compare our results with a previous co-training approach. To alleviate the problem of data sparseness, we create extra pseudo-examples for SCL by making queries to a search engine. Experiments on real-world on-line review data demonstrate the two techniques can effectively improve the performance compared to previous work.

Using Anaphora Resolution to Improve Opinion Target Identification in Movie Reviews
Niklas Jakob and Iryna Gurevych
Current work on automatic opinion mining has ignored opinion targets expressed by anaphorical pronouns, thereby missing a significant number of opinion targets. In this paper we empirically evaluate whether using an off-the-shelf anaphora resolution algorithm can improve the performance of a baseline opinion mining system. We present an analysis based on two different anaphora resolution systems. Our experiments on a movie review corpus demonstrate, that an unsupervised anaphora resolution algorithm significantly improves the opinion target extraction. We furthermore suggest domain and task specific extensions to an off-the-shelf algorithm which in turn yield significant improvements.

Hierarchical Sequential Learning for Extracting Opinions and Their Attributes
Yejin Choi and Claire Cardie
Automatic opinion recognition involves a number of related tasks, such as identifying the boundaries of opinion expression, determining their polarity, and determining their intensity. Although much progress has been made in this area, existing research typically treats each of the above tasks in isolation. In this paper, we apply a hierarchical parameter sharing technique using Conditional Random Fields for fine-grained opinion analysis, jointly detecting the boundaries of opinion expressions as well as determining two of their key attributes – polarity and intensity. Our experimental results show that our proposed approach improves the performance over a baseline that does not exploit hierarchical structure among the classes. In addition, we find that the joint approach outperforms a baseline that is based on cascading two separate components.

Short Talks: Morphology and Information Extraction, 11:55–13:15, Venue B, Lecture Hall 3

A Hybrid Rule/Model-Based Finite-State Framework for Normalizing SMS Messages
Richard Beaufort, Sophie Roekhaut, Louise-Amélie Cougnon and Cédrick Fairon
In recent years, research in natural language processing focused more and more on normalizing text messages. Several approaches were proposed, based either on standard spelling correction technics, translation models or speech recognition methods. However, the problem remains far from being solved: best systems achieve an accuracy of at best 60% at the sentence level, with a word error rate of at least 10%. In this paper, we present a hybrid approach, which combines both linguistics and statistics. The system involves four steps: a rule-based preprocessing, which splits the text into labeled units, like URLs or phones, and unlabeled parts, potentially noisy; a normalization step, relying on statistical models and exclusively performed on the unlabeled parts of the text; a morphosyntactic analysis of the normalized text; finally, a print step, which observes typography rules to build correct sentences, guided by the pieces of information provided by the linguistic analysis. The whole system, based on weighted finite-state machines, is part and parcel of a text-to-speech synthesis system.

Letter-Phoneme Alignment: An Exploration
Sittichai Jiampojamarn and Grzegorz Kondrak
Letter-phoneme alignment is usually generated by a straightforward application of the EM algorithm. We explore several alternative alignment methods that employ phonetics, integer programming, and sets of constraints, and propose a novel approach of refining the EM alignment by aggregation of best alignments. We perform both intrinsic and extrinsic evaluation of the assortment of methods. We show that our proposed EM-Aggregation algorithm leads to the improvement of the state of the art in letter-to-phoneme conversion on several different data sets.

Jointly Optimizing a Two-Step Conditional Random Field Model for Machine Transliteration and Its Fast Decoding Algorithm
Dong Yang, Paul Dixon and Sadaoki Furui
This paper presents a joint optimization method of a two-step conditional random field (CRF) model for machine transliteration and a fast decoding algorithm for the proposed method. Our method lies in the category of direct orthographical mapping (DOP) between two languages without using any intermediate phonemic mapping. In the two-step CRF model, the first CRF segments an input word into chunks and the second one converts each chunk into one unit in the target language. In this paper, we propose a method to jointly optimize the two-step CRFs and also a fast algorithm to realize it. Our experiments show that the proposed method outperforms the well-known joint source channel model (JSCM) and our proposed fast algorithm decreases the decoding time significantly. Furthermore, combination of the proposed method and the JSCM gives further improvement, which outperforms state-of-the-art results in terms of top-1 accuracy.

Arabic Named Entity Recognition: Using Features Extracted from Noisy Data
Yassine Benajiba, Imed Zitouni, Mona Diab and Paolo Rosso
Building an accurate Named Entity Recognition (NER) system for languages with complex morphology is a challenging task. In this paper, we present research that explores the feature space using both gold and bootstrapped noisy features to build an improved highly accurate Arabic NER system. We bootstrap noisy features by projection from an Arabic-English parallel corpus that is automatically tagged with a baseline NER system. The feature space covers lexical, morphological, and syntactic features. The proposed approach yields an improvement of up to 1.64 F-measure (absolute).

Extracting Sequences from the Web
Anthony Fader, Stephen Soderland and Oren Etzioni
Classical Information Extraction (IE) systems fill slots in domain-specific frames. This paper reports on Seq , a novel open IE system that leverages a domain-independent frame to extract ordered sequences such as presidents of the United States or the most common causes of death in the U.S. Seq leverages regularities about sequences to extract a coherent set of sequences from Web text. Seq nearly doubles the area under the precision-recall curve compared to an extractor that does not exploit these regularities.

An Entity-Level Approach to Information Extraction
Aria Haghighi and Dan Klein
We present a generative model of template-filling in which coreference resolution and role assignment are jointly determined. Underlying template roles first generate abstract entities, which in turn generate concrete textual mentions. On the standard corporate acquisitions dataset, joint resolution in our entity-level model reduces error over a mention-level discriminative approach by up to 20%.

Using Document Level Cross-Event Inference to Improve Event Extraction
Shasha Liao and Ralph Grishman

Event extraction is a particularly challenging type of information extraction (IE). Most current event extraction systems rely on local information at the phrase or sentence level. However, this local context may be insufficient to resolve ambiguities in identifying particular types of events; information from a wider scope can serve to resolve some of these ambiguities. In this paper, we use document level information to improve the performance of ACE event extraction. In contrast to previous work, we do not limit ourselves to information about events of the same type, but rather use information about other types of events to make predictions or resolve ambiguities regarding a given event. We learn such relationships from the training corpus and use them to help predict the occurrence of events and event arguments in a text. Experiments show that we can get 9.0% (absolute) gain in trigger (event) classification, and more than 8% gain for argument (role) classification in ACE event extraction.

A Semi-Supervised Key Phrase Extraction Approach: Learning from Title Phrases through a Document Semantic Network

Decong Li, Sujian Li, Wenjie Li, Wei Wang and Weiguang Qu

It is a fundamental and important task to extract key phrases from documents. Generally, phrases in a document are not independent in delivering the content of the document. In order to capture and make better use of their relationships in key phrase extraction, we suggest exploring the Wikipedia knowledge to model a document as a semantic network, where both n-ary and binary relationships among phrases are formulated. Based on a commonly accepted assumption that the title of a document is always elaborated to reflect the content of a document and consequently key phrases tend to have close semantics to the title, we propose a novel semi-supervised key phrase extraction approach in this paper by computing the phrase importance in the semantic network, through which the influence of title phrases is propagated to the other phrases iteratively. Experimental results demonstrate the remarkable performance of this approach.

Short Talks: Speech, Multimodal, and Summarization, 11:55–13:15, Venue B, Lecture Hall 4

Domain Adaptation of Maximum Entropy Language Models

Tanel Alumäe and Mikko Kurimo

We investigate a recently proposed Bayesian adaptation method for building style-adapted maximum entropy language models for speech recognition, given a large corpus of written language data and a small corpus of speech transcripts. Experiments show that the method consistently outperforms linear interpolation which is typically used in such cases.

Decision Detection Using Hierarchical Graphical Models

Trung H. Bui and Stanley Peters

We investigate hierarchical graphical models (HGMs) for automatically detecting decisions in multi-party discussions. Several types of dialogue act (DA) are distinguished on the basis of their roles in formulating decisions. HGMs enable us to model dependencies between observed features of discussions, decision DAs, and subdialogues that result in a decision. For the task of detecting decision regions, an HGM classifier outperforms non-hierarchical graphical models and support vector machines, raising the F1-score to 0.80 from 0.55.

Now, Where Was I? Resumption Strategies for an In-Vehicle Dialogue System

Jessica Villing

In-vehicle dialogue systems often contain more than one application, e.g. a navigation and a telephone application. This means that the user might, for example, interrupt the interaction with the telephone application to ask for directions from the navigation application, and then resume the dialogue with the telephone application. In this paper we present an analysis of interruption and resumption behaviour in human-human in-vehicle dialogues and also propose some implications

for resumption strategies in an in-vehicle dialogue system.

Using Speech to Reply to SMS Messages While Driving: An In-Car Simulator User Study
Yun-Cheng Ju and Tim Paek
Speech recognition affords automobile drivers a hands-free, eyes-free method of replying to Short Message Service (SMS) text messages. Although a voice search approach based on template matching has been shown to be more robust to the challenging acoustic environment of automobiles than using dictation, users may have difficulties verifying whether SMS response templates match their intended meaning, especially while driving. Using a high-fidelity driving simulator, we compared dictation for SMS replies versus voice search in increasingly difficult driving conditions. Although the two approaches did not differ in terms of driving performance measures, users made about six times more errors on average using dictation than voice search.

Learning to Follow Navigational Directions
Adam Vogel and Daniel Jurafsky
We present a system that learns to follow navigational natural language directions. Where traditional models learn from linguistic annotation or word distributions, our approach is grounded in the world, learning only from routes through a map paired with English descriptions. Lacking an explicit alignment between the text and the reference path makes it difficult to determine what portions of the language describe which aspects of the route. We learn this correspondence with a reinforcement learning algorithm, using the deviation of the route we follow from the intended path as a reward signal. We demonstrate that our system successfully grounds the meaning of spatial terms like 'above' and 'south' into geometric properties of paths.

Classification of Feedback Expressions in Multimodal Data
Costanza Navarretta and Patrizia Paggio
This paper addresses the issue of how linguistic feedback expressions, prosody and head gestures, i.e. head movements and face expressions, relate to one another in a collection of eight video-recorded Danish map-task dialogues. The study shows that in these data, prosodic features and head gestures significantly improve automatic classification of dialogue act labels for linguistic expressions of feedback.

A Hybrid Hierarchical Model for Multi-Document Summarization
Asli Celikyilmaz and Dilek Hakkani-Tur
Scoring sentences in documents given abstract summaries created by humans is important in extractive multi-document summarization. In this paper, we formulate extractive summarization as a two step learning problem building a generative model for pattern discovery and a regression model for inference. We calculate scores for sentences in document clusters based on their latent characteristics using a hierarchical topic model. Then, using these scores, we train a regression model based on the lexical and structural characteristics of the sentences, and use the model to score sentences of new documents to form a summary. Our system advances current state-of-the-art improving ROUGE scores by 7%. Generated summaries are less redundant and more coherent based upon manual quality evaluations.

Optimizing Informativeness and Readability for Sentiment Summarization
Hitoshi Nishikawa, Takaaki Hasegawa, Yoshihiro Matsuo and Genichiro Kikui
We propose a novel algorithm for sentiment summarization that takes account of informativeness and readability, simultaneously. Our algorithm generates a summary by selecting and ordering sentences taken from multiple review texts according to two scores that represent the informativeness and readability of the sentence order. The informativeness score is defined by the number of sentiment expressions and the readability score is learned from the target corpus. We evaluate our method by summarizing reviews on restaurants. Our method outperforms an existing

algorithm as indicated by its ROUGE score and human readability experiments.

Student Research Workshop Poster Session, 11:55–13:15, Venue A, Room VIII

Mood Patterns and Affective Lexicon Access in Weblogs
Thin Nguyen
The emergence of social media brings chances, but also challenges, to linguistic analysis. In this paper we investigate a novel problem of discovering patterns based on emotion and the association of moods and affective lexicon usage in blogosphere, a representative for social media. We propose the use of normative emotional scores for English words in combination with a psychological model of emotion measurement and a nonparametric clustering process for inferring meaningful emotion patterns automatically from data. Our results on a dataset consisting of more than 17 million mood-groundtruthed blogposts have shown interesting evidence of the emotion patterns automatically discovered that match well with the core-affect emotion model theorized by psychologists. We then present a method based on information theory to discover the association of moods and affective lexicon usage in the new media.

Growing Related Words from Seed via User Behaviors: A Re-ranking Based Approach
Yabin Zheng, Zhiyuan Liu and Lixing Xie
Motivated by Google Sets, we study the problem of growing related words from a single seed word by leveraging user behaviors hiding in user records of Chinese input method. Our proposed method is motivated by the observation that the more frequently two words co-occur in user records, the more related they are. First, we utilize user behaviors to generate candidate words. Then, we utilize search engine to enrich candidate words with adequate semantic features. Finally, we reorder candidate words according to their semantic relatedness to the seed word. Experimental results on a Chinese input method dataset show that our method gains better performance.

Transition-Based Parsing with Confidence-Weighted Classification
Martin Haulrich
We show that using confidence-weighted classification in transition-based parsing gives results comparable to using SVMs with faster training and parsing time. We also compare with other online learning algorithms and investigate the effect of pruning features when using confidence-weighted classification.

Expanding Verb Coverage in Cyc With VerbNet
Clifton McFate
A robust dictionary of semantic frames is an essential element of natural language understanding systems that use ontologies. However, creating lexical resources that accurately capture semantic representations en masse is a persistent problem. Where the sheer amount of content makes hand creation inefficient, computerized approaches often suffer from over generality and difficulty with sense disambiguation. This paper describes a semi-automatic method to create verb semantic frames in the Cyc ontology by converting the information contained in VerbNet into a Cyc usable format. This method captures the differences in meaning between types of verbs, and uses existing connections between WordNet, VerbNet, and Cyc to specify distinctions between individual verbs when available. This method provides 27,909 frames to OpenCyc which currently has none and can be used to extend ResearchCyc as well. We show that these frames lead to a 20% increase in sample sentences parsed over the Research Cyc verb lexicon.

A Framework for Figurative Language Detection Based on Sense Differentiation
Daria Bogdanova
Various text mining algorithms require the process of feature selection. High-level semantically rich features, such as figurative language uses, speech errors etc., are very promising for such

problems as e.g. writing style detection, but automatic extraction of such features is a big challenge. In this paper, we propose a framework for figurative language use detection. This framework is based on the idea of sense differentiation. We describe two algorithms illustrating the mentioned idea. We show then how these algorithms work by applying them to Russian language data.

Automatic Selectional Preference Acquisition for Latin verbs
Barbara McGillivray
We present a system that automatically induces Selectional Preferences (SPs) for Latin verbs from two treebanks by using Latin WordNet. Our method overcomes some of the problems connected with data sparseness and the small size of the input corpora. We also suggest a way to evaluate the acquired SPs on unseen events extracted from other Latin corpora.

Edit Tree Distance Alignments for Semantic Role Labelling
Hector-Hugo Franco-Penya
"Tree SRL system" is a Semantic Role Label-ling supervised system based on a tree-distance algorithm and a simple k-NN implementation. The novelty of the system lies in comparing the sentences as tree structures with multiple rela-tions, instead of extracting vectors of features for each relation and classifying them. The sys-tem was tested with the English CoNLL-2009 shared task data set where 79% accuracy was obtained.

Automatic Sanskrit Segmentizer Using Finite State Transducers
Vipul Mittal
In this paper, we propose a novel method for automatic segmentation of a Sanskrit string into different words. The input for our segmentizer is a Sanskrit string either encoded as a Unicode string or as a Roman transliterated string and the output is a set of possible splits with saliency associated with each of them. We followed two different approaches to segment a Sanskrit text using sandhi rules extracted from a parallel corpus of manually sandhi split text. While the first approach augments the finite state transducer used to analyze Sanskrit morphology and traverse it to segment a word, the second approach generates all possible segmentations and validates each constituent using a morph analyzer.

Adapting Self-training for Semantic Role Labeling
Rasoul Samad Zadeh Kaljahi
Supervised semantic role labeling (SRL) systems trained on hand-crafted annotated corpora have recently achieved state-of-the-art performance. However, creating such corpora is tedious and costly, with the resulting corpora not sufficiently representative of the language. This paper describes a part of an ongoing work on applying bootstrapping methods to SRL to deal with this problem. Previous work shows that, due to the complexity of SRL, this task is not straight forward. One major difficulty is the propagation of classification noise into the successive iterations. We address this problem by employing balancing and preselection methods for self-training, as a bootstrapping algorithm. The proposed methods could achieve improvement over the base line, which do not use these methods.

Weakly Supervised Learning of Presupposition Relations between Verbs
Galina Tremper
Presupposition relations between verbs are not very well covered in existing lexical semantic resources. We propose a weakly supervised algorithm for learning presupposition relations between verbs that distinguishes five semantic relations: presupposition, entailment, temporal inclusion, antonymy and other/no relation. We start with a number of seed verb pairs selected manually for each semantic relation and classify unseen verb pairs. Our algorithm achieves an overall accuracy of 36% for type-based classification.

Importance of Linguistic Constraints in Statistical Dependency Parsing

Bharat Ram Ambati
Statistical systems with high accuracy are very useful in real-world applications. If these systems can capture basic linguistic information, then the usefulness of these statistical systems improve a lot. This paper is an attempt at incorporating linguistic constraints in statistical dependency parsing. We consider a simple linguistic constraint that a verb should not have multiple subjects/objects as its children in the dependency tree. We first describe the importance of this constraint considering Machine Translation systems which use dependency parser output, as an example application. We then show how the current state-of-the-art dependency parsers violate this constraint. We present two new methods to handle this constraint. We evaluate our methods on the state-of-the-art dependency parsers for Hindi and Czech.

The Use of Formal Language Models in the Typology of the Morphology of Amerindian Languages
Andres Osvaldo Porta
The aim of this work is to present some preliminary results of an investigation in course on the typology of the morphology of the native South American languages from the point of view of the formal language theory. With this object, we give two contrasting examples of descriptions of two Aboriginal languages finite verb forms morphology: Argentinean Quechua (quichua santiagueño) and Toba. The description of the morphology of the finite verb forms of Argentinean Quechua uses finite automata and finite transducers. In this case the construction is straightforward using two level morphology and then, describes in a very natural way the Argentinean Quechua morphology using a regular language. On the contrary, the Toba verbs morphology, with a system that uses simultaneously prefixes and suffixes, has not a natural description as regular language. Toba has a complex system of causative suffixes, whose successive applications determinate the use of prefixes belonging different person marking prefix sets. We adopt the solution of Creider et al. (1995) to naturally deal with this and other similar morphological processes which involve interactions between prefixes and suffixes and then we describe the toba morphology using linear context-free languages and two-taped automata.

Poster Session 2, 13:15–15:00, Venue A, Foyer

Last but Definitely Not Least: On the Role of the Last Sentence in Automatic Polarity-Classification
Israela Becker and Vered Aharonson
Two psycholinguistic and psychophysical experiments show that in order to efficiently extract polarity of written texts such as customer reviews on the Internet, one should concentrate computational efforts on messages in the final position of the text.

Automatically Generating Annotator Rationales to Improve Sentiment Classification
Ainur Yessenalina, Yejin Choi and Claire Cardie
One of the central challenges in sentiment-based text categorization is that not every portion of a document is equally informative for inferring the overall sentiment of the document. Previous research has shown that enriching the sentiment labels with human annotators' "rationales" can produce substantial improvements in categorization performance (Zaidan et al., 2007). We explore methods to automatically generate annotator rationales for document-level sentiment classification. Rather unexpectedly, we find the automatically generated rationales just as helpful as human rationales.

Simultaneous Tokenization and Part-Of-Speech Tagging for Arabic without a Morphological Analyzer
Seth Kulick
We describe an approach to simultaneous tokenization and part-of-speech tagging that is based

on separating the closed and open-class items, and focusing on the likelihood of the possible stems of the open-class words. By encoding some basic linguistic information, the machine learning task is simplified, while achieving state-of-the-art tokenization results and competitive POS results, although with a reduced tag set and some evaluation difficulties.

Hierarchical A* Parsing with Bridge Outside Scores
Adam Pauls and Dan Klein
Hierarchical A* (HA*) uses of a hierarchy of coarse grammars to speed up parsing without sacrificing optimality. HA* prioritizes search in refined grammars using Viterbi outside costs computed in coarser grammars. We present Bridge Hierarchical A* (BHA*), a modified Hierarchial A* algorithm which computes a novel outside cost called a *bridge* outside cost. These bridge costs mix finer outside scores with coarser inside scores, and thus constitute tighter heuristics than entirely coarse scores. We show that BHA* substantially outperforms HA* when the hierarchy contains only very coarse grammars, while achieving comparable performance on more refined hierarchies.

Using Parse Features for Preposition Selection and Error Detection
Joel Tetreault, Jennifer Foster and Martin Chodorow
We evaluate the effect of adding parse features to a leading model of preposition usage. Results show a significant improvement in the preposition selection task on native speaker text and modest increments in precision and recall in an ESL error detection task. Analysis of the parser output indicates that it is robust enough in the face of noisy non-native writing to extract useful information.

Distributional Similarity vs. PU Learning for Entity Set Expansion
Xiao-Li Li, Lei Zhang, Bing Liu and See-Kiong Ng
Distributional similarity is a classic tech-nique for entity set expansion, where the system is given a set of seed entities of a particular class, and is asked to expand the set using a corpus to obtain more entities of the same class as represented by the seeds. This paper shows that a machine learning model called positive and unla-beled learning (PU learning) can model the set expansion problem better. Based on the test results of 10 corpora, we show that a PU learning technique outperformed distributional similarity significantly.

Active Learning-Based Elicitation for Semi-Supervised Word Alignment
Vamshi Ambati, Stephan Vogel and Jaime Carbonell
Semi-supervised word alignment aims to improve the accuracy of automatic word alignment by incorporating full or partial manual alignments. Motivated by standard active learning query sampling frameworks such as uncertainty-, margin- and query-by-committee sampling we propose multiple query strategies for the alignment link selection task. Our experiments show that by active selection of uncertain and informative links, we reduce the overall manual effort involved in elicitation of alignment link data for training a semi-supervised word aligner.

An Active Learning Approach to Finding Related Terms
David Vickrey, Oscar Kipersztok and Daphne Koller
We present a novel system that helps non-experts find sets of similar words. The user begins by specifying one or more seed words. The system then iteratively suggests a series of candidate words, which the user can either accept or reject. Current techniques for this task typically bootstrap a classifier based on a fixed seed set. In contrast, our system involves the user throughout the labeling process, using active learning to intelligently explore the space of similar words. In particular, our system can take advantage of negative examples provided by the user. Our system combines multiple pre-existing sources of similarity data (a standard thesaurus, WordNet, contextual similarity), enabling it to capture many types of similarity groups ("synonyms of crash," "types of car," etc.). We evaluate on a hand-labeled evaluation set; our system improves over a strong baseline by 36%.

Learning Better Data Representation Using Inference-Driven Metric Learning
Paramveer S. Dhillon, Partha Pratim Talukdar and Koby Crammer
We initiate a study comparing effectiveness of the transformed spaces learned by recently proposed supervised, and semi-supervised metric learning algorithms to those generated by previously proposed unsupervised dimensionality reduction methods (e.g., PCA). Through a variety of experiments on different real-world datasets, we find IDML-IT, a semi-supervised metric learning algorithm to be the most effective.

Wrapping up a Summary: From Representation to Generation
Josef Steinberger, Marco Turchi, Mijail Kabadjov, Ralf Steinberger and Nello Cristianini
The main focus of this work is to investigate robust ways for generating summaries from summary representations without recurring to simple sentence extraction and aiming at more human-like summaries. This is motivated by empirical evidence from TAC 2009 data showing that human summaries contain on average more and shorter sentences than the system summaries. We report encouraging preliminary results comparable to those attained by participating systems at TAC 2009.

Translation and Multilinguality, 15:00–16:15, Venue A, Aula

Improving Statistical Machine Translation with Monolingual Collocation
Zhanyi Liu, Haifeng Wang, Hua Wu and Sheng Li
This paper proposes to use monolingual collocations to improve Statistical Machine Translation (SMT). We make use of the collocation probabilities, which are estimated from monolingual corpora, in two aspects, namely improving word alignment for various kinds of SMT systems and improving phrase table for phrase-based SMT. The experimental results show that our method improves the performance of both word alignment and translation quality significantly. As compared to baseline systems, we achieve absolute improvements of 2.40 BLEU score on a phrase-based SMT system and 1.76 BLEU score on a parsing-based SMT system.

Bilingual Sense Similarity for Statistical Machine Translation
Boxing Chen, George Foster and Roland Kuhn
This paper proposes new algorithms to compute the sense similarity between two units (words, phrases, rules, etc.) from parallel corpora. The sense similarity scores are computed by using the vector space model. We then apply the algorithms to statistical machine translation by computing the sense similarity be-tween the source and target side of translation rule pairs. Similarity scores are used as additional features of the translation model to improve translation performance. Significant improvements are obtained over a state-of-the-art hierarchical phrase-based machine translation system.

Untangling the Cross-Lingual Link Structure of Wikipedia
Gerard de Melo and Gerhard Weikum
Wikipedia articles in different languages are connected by interwiki links that are increasingly being recognized as a valuable source of cross-lingual information. Unfortunately, large numbers of links are imprecise or simply wrong. In this paper, techniques to detect such problems are identified. We formalize their removal as an optimization task based on graph repair operations. We then present an algorithm with provable properties that uses linear programming and a region growing technique to tackle this challenge. This allows us to transform Wikipedia into a much more consistent multilingual register of the world's entities and concepts.

Machine Learning, 15:00–16:15, Venue A, Hall X

Bucking the Trend: Large-Scale Cost-Focused Active Learning for Statistical Machine Transla-

tion
Michael Bloodgood and Chris Callison-Burch
We explore how to improve machine translation systems by adding more translation data in situations where we already have substantial resources. The main challenge is how to buck the trend of diminishing returns that is commonly encountered. We present an active learning-style data solicitation algorithm to meet this challenge. We test it, gathering annotations via Amazon Mechanical Turk, and find that we get an order of magnitude increase in performance rates of improvement.

Creating Robust Supervised Classifiers via Web-Scale N-Gram Data
Shane Bergsma, Emily Pitler and Dekang Lin
In this paper, we systematically assess the value of using web-scale N-gram data in state-of-the-art supervised NLP classifiers. We compare classifiers that include or exclude features for the counts of various N-grams, where the counts are obtained from a web-scale auxiliary corpus. We show that including N-gram count features can advance the state-of-the-art accuracy on standard data sets for adjective ordering, spelling correction, noun compound bracketing, and verb part-of-speech disambiguation. More importantly, when operating on new domains, or when labeled training data is not plentiful, we show that using web-scale N-gram features is essential for achieving robust performance.

Convolution Kernel over Packed Parse Forest
Min Zhang, Hui Zhang and Haizhou Li
This paper proposes a convolution forest kernel to effectively explore rich structured features embedded in a packed parse forest. As opposed to the convolution tree kernel, the proposed forest kernel does not have to commit to a single best parse tree, is thus able to explore very large object spaces and much more structured features embedded in a forest. This makes the proposed kernel more robust against parsing errors and data sparseness issues than the convolution tree kernel. The paper presents the formal definition of convolution forest kernel and also illustrates the computing algorithm to fast compute the proposed convolution forest kernel. Experimental results on two NLP applications, relation extraction and semantic role labeling, show that the proposed forest kernel significantly outperforms the baseline of the convolution tree kernel.

Language Learning and Models of Language, 15:00–16:15, Venue A, Hall IX

Estimating Strictly Piecewise Distributions
Jeffrey Heinz and James Rogers
Strictly Piecewise (SP) languages are a subclass of regular languages which encode certain kinds of long-distance dependencies that are found in natural languages. Like the classes in the Chomsky and Subregular hierarchies, there are many independently converging characterizations of the SP class (Rogers et al., to appear). Here we define SP distributions and show that they can be efficiently estimated from positive data.

String Extension Learning
Jeffrey Heinz
This paper provides a unified, learning-theoretic analysis of several learnable classes of languages discussed previously in the literature. The analysis shows that for these classes an incremental, globally consistent, locally conservative, set-driven learner always exists. Additionally, the analysis provides a recipe for constructing new learnable classes. Potential applications include learnable models for aspects of natural language and cognition.

Compositional Matrix-Space Models of Language
Sebastian Rudolph and Eugenie Giesbrecht
We propose CMSMs, a novel type of generic compositional models for syntactic and semantic

aspects of natural language, based on matrix multiplication. We argue for the structural and cognitive plausibility of this model and show that it is able to cover and combine various common compositional NLP approaches ranging from statistical word space models to symbolic grammar formalisms.

Summarization 2, 15:00–16:15, Venue B, Lecture Hall 3

Cross-Language Document Summarization Based on Machine Translation Quality Prediction
Xiaojun Wan, Huiying Li and Jianguo Xiao
Cross-language document summarization is a task of producing a summary in one language for a document set in a different language. Existing methods simply use machine translation for document translation or summary translation. However, current machine translation services are far from satisfactory, which results in that the quality of the cross-language summary is usually very poor, both in readability and content. In this paper, we propose to consider the translation quality of each sentence in the English-to-Chinese cross-language summarization process. First, the translation quality of each English sentence in the document set is predicted with the SVM regression method, and then the quality score of each sentence is incorporated into the summarization process. Finally, the English sentences with high translation quality and high informativeness are selected and translated to form the Chinese summary. Experimental results demonstrate the effectiveness and usefulness of the proposed approach.

A New Approach to Improving Multilingual Summarization Using a Genetic Algorithm
Marina Litvak, Mark Last and Menahem Friedman
Automated summarization methods can be defined as "language-independent," if they are not based on any languagespecific knowledge. Such methods can be used for multilingual summarization defined by Mani (2001) as "processing several languages, with summary in the same language as input." In this paper, we introduce MUSE, a languageindependent approach for extractive summarization based on the linear optimization of several sentence ranking measures using a genetic algorithm. We tested our methodology on two languages—English and Hebrew—and evaluated its performance with ROUGE-1 Recall vs. stateof- the-art extractive summarization approaches. Our results show that MUSE performs better than the best known multilingual approach (TextRank) in both languages. Moreover, our experimental results on a bilingual (English and Hebrew) document collection suggest that MUSE does not need to be retrained on each language and the same model can be used across at least two different languages.

Bayesian Synchronous Tree-Substitution Grammar Induction and Its Application to Sentence Compression
Elif Yamangil and Stuart M. Shieber
We describe our experiments with training algorithms for tree-to-tree synchronous tree-substitution grammar (STSG) for monolingual translation tasks such as sentence compression and paraphrasing. These translation tasks are characterized by the relative ability to commit to parallel parse trees and availability of word alignments, yet the unavailability of large-scale data, calling for a Bayesian tree-to-tree formalism. We formalize nonparametric Bayesian STSG with epsilon alignment in full generality, and provide a Gibbs sampling algorithm for posterior inference tailored to the task of extractive sentence compression. We achieve improvements against a number of baselines, including expectation maximization and variational Bayes training, illustrating the merits of nonparametric inference over the space of grammars as opposed to sparse parametric inference with a fixed grammar.

Semantics 3, 15:00–16:15, Venue B, Lecture Hall 4

Contextualizing Semantic Representations Using Syntactically Enriched Vector Models

Stefan Thater, Hagen Fürstenau and Manfred Pinkal
We present a syntactically enriched vector model that supports the computation of contextualized semantic representations in a quasi compositional fashion. It employs a systematic combination of first- and second-order context vectors. We apply our model to two different tasks, and show that (i) it substantially outperforms previous work on a paraphrase ranking task, and (ii) achieves promising results on a word-sense similarity task; to our knowledge, it is the first time that an unsupervised method has been applied to this task.

Bootstrapping Semantic Analyzers from Non-Contradictory Texts
Ivan Titov and Mikhail Kozhevnikov
We argue that groups of unannotated texts with overlapping and non-contradictory semantics represent a valuable source of information for learning semantic representations. A simple and efficient inference method recursively induces joint semantic representations for each group and discovers correspondence between lexical entries and latent semantic concepts. We consider the generative semantics-text correspondence model (Liang et al.,2009) and demonstrate that exploiting the non-contradiction relation between texts leads to substantial improvements over natural baselines on a problem of analyzing human-written weather forecasts.

Open-Domain Semantic Role Labeling by Modeling Word Spans
Fei Huang and Alexander Yates
Most supervised language processing systems show a significant drop-off in performance when they are tested on text that comes from a domain significantly different from the domain of the training data. Semantic role labeling techniques are typically trained on newswire text, and in tests their performance on fiction is as much as 19% worse than their performance on newswire text. We investigate techniques for building open-domain semantic role labeling systems that approach the ideal of a train-once, use-anywhere system. We leverage recently-developed techniques for learning representations of text using latent-variable language models, and extend these techniques to ones that provide the kinds of features that are useful for semantic role labeling. In experiments, our novel system reduces error by 16% relative to the previous state of the art on out-of-domain text.

Software Demonstration Session, 15:00–17:35, Venue A, Room XI

Grammar Prototyping and Testing with the LinGO Grammar Matrix Customization System
Emily M. Bender, Scott Drellishak, Antske Fokkens, Michael Wayne Goodman, Daniel P. Mills, Laurie Poulson and Safiyyah Saleem
This demonstration presents the LinGO Grammar Matrix grammar customization system: a repository of distilled linguistic knowledge and a web-based service which elicits a typological description of a language from the user and yields a customized grammar fragment ready for sustained development into a broad-coverage grammar. We describe the implementation of this repository with an emphasis on how the information is made available to users, including in-browser testing capabilities.

cdec: A Decoder, Alignment, and Learning Framework for Finite-State and Context-Free Translation Models
Chris Dyer, Adam Lopez, Juri Ganitkevitch, Jonathan Weese, Ferhan Ture, Phil Blunsom, Hendra Setiawan, Vladimir Eidelman and Philip Resnik
We present cdec, an open source framework for decoding, aligning with, and training a number of statistical machine translation models, including word-based models, phrase-based models, and models based on synchronous context-free grammars. Using a single unified internal representation for translation forests, the decoder strictly separates model-specific translation logic from general rescoring, pruning, and inference algorithms. From this unified representation, the decoder

can extract not only the 1- or k-best translations, but also alignments to a reference, or the quantities necessary to drive discriminative training using gradient-based or gradient-free optimization techniques. Its efficient C++ implementation means that memory use and runtime performance are significantly better than comparable decoders.

Beetle II: A System for Tutoring and Computational Linguistics Experimentation
Myroslava O. Dzikovska, Johanna D. Moore, Natalie Steinhauser, Gwendolyn Campbell, Elaine Farrow and Charles B. Callaway
We present Beetle II, a tutorial dialogue system designed to accept unrestricted language input and support experimentation with different tutorial planning and dialogue strategies. Our first ystem evaluation used two different tutorial policies and demonstrated that the system can be successfully used to study the impact of different approaches to tutoring. In the future, the system can also be used to experiment with a variety of natural language interpretation and generation techniques.

GernEdiT – The GermaNet Editing Tool
Verena Henrich and Erhard Hinrichs
GernEdiT (short for: GermaNet Editing Tool) offers a graphical interface for the lexicographers and developers of GermaNet to access and modify the underlying GermaNet resource. GermaNet is a lexical-semantic wordnet that is modeled after the Princeton WordNet for English. The traditional lexicographic development of GermaNet was error prone and time-consuming, mainly due to a complex underlying data format and no opportunity of automatic consistency checks. GernEdiT replaces the earlier development by a more user-friendly tool, which facilitates automatic checking of internal consistency and correctness of the linguistic resource. This paper presents all these core functionalities of GernEdiT along with details about its usage and usability.

WebLicht: Web-Based LRT Services for German
Erhard Hinrichs, Marie Hinrichs and Thomas Zastrow
This software demonstration presents WebLicht (short for: Web-Based Linguistic Chaining Tool), a web-based service environment for the integration and use of language resources and tools (LRT). WebLicht is being developed as part of the D-SPIN project . We-bLicht is implemented as a web application so that there is no need for users to install any software on their own computers or to concern themselves with the technical details involved in building tool chains. The integrated web services are part of a prototypical infrastructure that was developed to facilitate chaining of LRT services. WebLicht allows the integration and use of distributed web services with standardized APIs. The nature of these open and standardized APIs makes it possible to access the web services from nearly any programming language, shell script or workflow engine (UIMA, Gate etc.) Additionally, an application for integration of additional services is available, allowing anyone to contribute his own web service.

The S-Space Package: An Open Source Package for Word Space Models
David Jurgens and Keith Stevens
We present the S-Space Package, an open source framework for developing and evaluating word space algorithms. The package implements well-known word space algorithms, such as LSA, and provides a comprehensive set of matrix utilities and data structures for extending new or existing models. The package also includes word space benchmarks for evaluation. Both algorithms and libraries are designed for high concurrency and scalability. We demonstrate the efficiency of the reference implementations and also provide their results on six benchmarks.

Talking NPCs in a Virtual Game World
Tina Klüwer, Peter Adolphs, Feiyu Xu, Hans Uszkoreit and Xiwen Cheng
The submission describes a system using dialog, information extraction and Semantic Web tech-

nologies to enable natural language for Non Player Characters (NPCs) in an online game world. Depending on the type of game, NPCs are often used for enhancing plot and challenges and for making the artificial world more vivid and therefore also more immersive. They can also help to populate new worlds by carrying out jobs the user-led characters come in touch with. The range of functions to be filled by NPCs is currently still strongly restricted by their limited capabilities in autonomous acting and communication. This shortcoming creates a strong need for progress in AI and NLP, especially in the areas of planning and dialogue systems.

An Open-Source Package for Recognizing Textual Entailment

Milen Kouylekov and Matteo Negri

This paper presents a general-purpose open source package for recognizing Textual Entailment. The system implements a collection of algorithms, providing a configurable framework to quickly set up a working environment to experiment with the RTE task. Fast prototyping of new solutions is also allowed by the possibility to extend its modular architecture. We present the tool as a useful resource to approach the Textual Entailment problem, as an instrument for didactic purposes, and as an opportunity to create a collaborative environment to promote research in the field.

Personalising Speech-To-Speech Translation in the EMIME Project

Mikko Kurimo, William Byrne, John Dines, Philip N. Garner, Matthew Gibson, Yong Guan, Teemu Hirsimäki, Reima Karhila, Simon King, Hui Liang, Keiichiro Oura, Lakshmi Saheer, Matt Shannon, Sayaki Shiota and Jilei Tian

In the EMIME project we have studied unsupervised cross-lingual speaker adaptation. We have employed an HMM statistical framework for both speech recognition and synthesis which provides transformation mechanisms to adapt the synthesized voice in TTS (text-to-speech) using the recognized voice in ASR (automatic speech recognition). An important application for this research is personalised speech-to-speech translation that will use the voice of the speaker in the input language to utter the translated sentences in the output language. In mobile environments this enhances the users' interaction across language barriers by making the output speech sound more like the original speaker's way of speaking, even if she or he could not speak the output language.

Hunting for the Black Swan: Risk Mining from Text

Jochen Leidner and Frank Schilder

In the business world, analyzing and dealing with risk permeates all decisions and actions. However, to date risk identification, the first step in the risk management cycle, has always been a manual activity with little to no intelligent software tool support. In addition, although companies are required to list risks to their business in their annual SEC filings in the USA, these descriptions are often very high-level and vague. In this paper, we introduce Risk Mining, which is the task of identifying a set of risks pertaining to a business area or entity. We argue that by combining Web mining and Information Extraction (IE) techniques, risks can be detected automatically before they materialize, thus providing valuable business intelligence. We describe a system that induces a risk taxonomy with concrete risks (e.g., interest rate changes) at its leaves and more abstract risks (e.g., financial risks) closer to its root node. The taxonomy is induced via a bootstrapping algorithms starting with a few seeds. The risk taxonomy is used by the system as input to a risk monitor that matches risk mentions in financial documents to the abstract risk types, thus bridging a lexical gap. Our system is able to automatically generate company specific "risk maps", which we demonstrate for a corpus of earnings report conference calls.

Speech-Driven Access to the Deep Web on Mobile Devices

Taniya Mishra and Srinivas Bangalore

The Deep Web is the collection of information repositories that are not indexed by search engines. These repositories are typically accessible through web forms and contain dynamically changing

information. In this paper, we present a system that allows users to access such rich repositories of information on mobile devices using spoken language.

Tools for Multilingual Grammar-Based Translation on the Web
Aarne Ranta, Krasimir Angelov and Thomas Hallgren
This is a system demo for a set of tools for translating texts between multiple languages in real time with high quality. The translation works on restricted languages, and is based on semantic interlinguas. The underlying model is GF (Grammatical Framework), which is an open-source toolkit for multilingual grammar implementations. The demo will cover up to 20 parallel languages. Two related sets of tools are presented: grammarian's tools helping to build translators for new domains and languages, and translator's tools helping to translate documents. The grammarian's tools are designed to make it easy to port the technique to new applications. The translator's tools are essential in the restricted language context, enabling the author to remain in the fragments recognized by the system. The tools that are demonstrated will be applied and developed further in the European project MOLTO (Multilingual On-Line Translation. FP7-ICT-247914), which will start in March 2010 and run for three years.

Demonstration of a Prototype for a Conversational Companion for Reminiscing about Images
Yorick Wilks, Roberta Catizone, Alexiei Dingli and Weiwei Cheng
This paper describes an initial prototype demonstrator of a Companion, designed as a platform for novel approaches to the following: 1) The use of Information Extraction (IE) techniques to extract the content of incoming dialogue utterances after an Automatic Speech Recognition (ASR) phase, 2) The conversion of the input to Resource Descriptor Format (RDF) to allow the generation of new facts from existing ones, under the control of a Dialogue Manger (DM), that also has access to stored knowledge and to open knowledge accessed in real time from the web, all in RDF form, 3) A DM implemented as a stack and network virtual machine that models mixed initiative in dialogue control, and 4) A tuned dialogue act detector based on corpus evidence. The prototype platform was evaluated, and we describe this briefly; it is also designed to support more extensive forms of emotion detection carried by both speech and lexical content, as well as extended forms of machine learning.

It Makes Sense: A Wide-Coverage Word Sense Disambiguation System for Free Text
Zhi Zhong and Hwee Tou Ng
Word sense disambiguation (WSD) systems based on supervised learning achieved the best performance in SensEval and SemEval workshops. However, there are few publicly available open source WSD systems. This limits the use of WSD in other applications, especially for researchers whose research interests are not in WSD. In this paper, we present IMS, a supervised English all-words WSD system. The flexible framework of IMS allows users to integrate different preprocessing tools, additional features, and different classifiers. By default, we use linear support vector machines as the classifier with multiple knowledge-based features. In our implementation, IMS achieves state-of-the-art results on several SensEval and SemEval tasks.

Semantics 4, 16:45–17:35, Venue A, Aula

Learning Script Knowledge with Web Experiments
Michaela Regneri, Alexander Koller and Manfred Pinkal
We describe a novel approach to unsupervised learning of the events that make up a script, along with constraints on their temporal ordering. We collect natural-language descriptions of script-specific event sequences from volunteers over the Internet. Then we compute a graph representation of the script's temporal structure using a multiple sequence alignment algorithm. The evaluation of our system shows that we outperform two informed baselines.

Starting from Scratch in Semantic Role Labeling
Michael Connor, Yael Gertner, Cynthia Fisher and Dan Roth
A fundamental step in sentence comprehension involves assigning semantic roles to sentence constituents. To accomplish this, the listener must parse the sentence, find constituents that are candidate arguments, and assign semantic roles to those constituents. Each step depends on prior lexical and syntactic knowledge. Where do children learning their first languages begin in solving this problem? In this paper we focus on the parsing and argument-identification steps that precede Semantic Role Labeling (SRL) training. We combine a simplified SRL with an unsupervised HMM part of speech tagger, and experiment with psycholinguistically-motivated ways to label clusters resulting from the HMM so that they can be used to parse input for the SRL system. The results show that proposed shallow representations of sentence structure are robust to reductions in parsing accuracy, and that the contribution of alternative representations of sentence structure to successful semantic role labeling varies with the integrity of the parsing and argument-identification stages.

Dialogue, 16:45–17:35, Venue A, Hall X

Modeling Norms of Turn-Taking in Multi-Party Conversation
Kornel Laskowski
Substantial research effort has been invested in recent decades into the computational study and automatic processing of multi-party conversation. While most aspects of conversational speech have benefited from a wide availability of analytic, computationally tractable techniques, only qualitative assessments are available for characterizing multi-party turn-taking. The current paper attempts to address this deficiency by first proposing a framework for computing turn-taking model perplexity, and then by evaluating several multi-participant modeling approaches. Experiments show that direct multi-participant models do not generalize to held out data, and likely never will, for practical reasons. In contrast, the Extended-Degree-of-Overlap model represents a suitable candidate for future work in this area, and is shown to successfully predict the distribution of speech in time and across participants in previously unseen conversations.

Optimising Information Presentation for Spoken Dialogue Systems
Verena Rieser, Oliver Lemon and Xingkun Liu
We present a novel approach to Information Presentation (IP) in Spoken Dialogue Systems (SDS) using a data-driven statistical optimisation framework for content planning and attribute selection. First we collect data in a Wizard-of-Oz (WoZ) experiment and use it to build a supervised model of human behaviour. This forms a baseline for measuring the performance of optimised policies, developed from this data using Reinforcement Learning (RL) methods. We show that the optimised policies significantly outperform the baselines in a variety of generation scenarios: while the supervised model is able to attain up to 87.6% of the possible reward on this task, the RL policies are significantly better in 5 out of 6 scenarios, gaining up to 91.5% of the total possible reward. The RL policies perform especially well in more complex scenarios. We are also the first to show that adding predictive "lower level" features (e.g. from the NLG realiser) is important for optimising IP strategies according to user preferences. This provides new insights into the nature of the IP problem for SDS.

Historical Linguistics, 16:45–17:35, Venue A, Hall IX

Combining Data and Mathematical Models of Language Change
Morgan Sonderegger and Partha Niyogi
English noun/verb (N/V) pairs ("contract", "cement") have undergone complex patterns of change between 3 stress patterns for several centuries. We describe a longitudinal dataset of N/V pair

pronunciations, leading to a set of properties to be accounted for by any computational model. We analyze the dynamics of 5 dynamical systems models of linguistic populations, each derived from a model of learning by individuals. We compare each model's dynamics to a set of properties observed in the N/V data, and reason about how assumptions about individual learning affect population-level dynamics.

Finding Cognate Groups Using Phylogenies
David Hall and Dan Klein
A central problem in historical linguistics is the identification of historically related cognate words. We present a generative phylogenetic model for automatically inducing cognate group structure from unaligned word lists. Our model represents the process of transformation and transmission from ancestor word to daughter word, as well as the alignment between the words lists of the observed languages. We also present a novel method for simplifying complex weighted automata created during inference to counteract the otherwise exponential growth of message sizes. On the task of identifying cognates in a dataset of Romance words, our model significantly outperforms a baseline approach, increasing accuracy by as much as 80%. Finally, we demonstrate that our automatically induced groups can be used to successfully reconstruct ancestral words.

Decipherment, 16:45–17:35, Venue B, Lecture Hall 3

An Exact A* Method for Deciphering Letter-Substitution Ciphers
Eric Corlett and Gerald Penn
This paper presents an algorithm for decoding monalphabetic ciphers, with the purpose in mind of automatically learning nonstandard encodings of electronic documents in which the language in known. This is useful in languages such as Hindi in which there is no dominant electronic standard for encoding the writing system. We present a set of tests for our algorithm and find that it gives highly accurate results, and that it has the potential to achieve very good running times.

A Statistical Model for Lost Language Decipherment
Benjamin Snyder, Regina Barzilay and Kevin Knight
In this paper we propose a method for the automatic decipherment of lost langauges. Given a non-parallel corpus in a known related language, our model produces both alphabetic mappings and translations of words into their corresponding cognates. We employ a non-parametric Bayesian framework to simultaneously capture both low-level character mappings and high-level morphemic correspondences. This formulation enables us to encode some of the linguistic intuitions that have guided human decipherers. When applied to the ancient Semitic language Ugaritic, the model correctly maps nearly all letters to their Hebrew counterparts, and deduces the correct Hebrew cognate for over half of the Ugaritic words which have cognates in Hebrew.

Tree Transducers, 16:45–17:35, Venue B, Lecture Hall 4

Efficient Inference through Cascades of Weighted Tree Transducers
Jonathan May, Kevin Knight and Heiko Vogler
Weighted tree transducers have been proposed as useful formal models for representing syntactic natural language processing applications, but there has been little description of inference algorithms for these automata beyond formal foundations. We give a detailed description of algorithms for application of cascades of weighted tree transducers to weighted tree acceptors, connecting formal theory with actual practice. Additionally, we present novel on-the-fly variants of these algorithms, and compare their performance on a syntax machine translation cascade based on (Yamada and Knight, 2001).

A Tree Transducer Model for Synchronous Tree-Adjoining Grammars

Andreas Maletti
A characterization of the expressive power of synchronous tree-adjoining grammars (STAGs) in terms of tree transducers (or equivalently, synchronous tree substitution grammars) is developed. Essentially, a STAG corresponds to an extended tree transducer that uses explicit substitution in both the input and output. This characterization allows the easy integration of STAG into toolkits for extended tree transducers. Moreover, the applicability of the characterization to several representational and algorithmic problems is demonstrated.

ACL 2010 Main Conference Abstracts: Wednesday, July 14

Parsing 3, 10:30–12:10, Venue A, Aula

Dynamic Programming for Linear-Time Incremental Parsing
Liang Huang and Kenji Sagae
Incremental parsing techniques such as shift-reduce have gained popularity thanks to their efficiency, but there remains a major problem: the search is *greedy*, and only explores a tiny fraction of the whole space (even with beam search) as opposed to dynamic programming. We show that, surprisingly, dynamic programming is in fact possible for many shift-reduce parsers, by merging "equivalent" stacks based on feature values. Empirically, our algorithm yields up to a five-fold speedup against conventional beam-search over a state-pf-the-art shift-reduce dependency parser with no loss in accuracy. Better search also leads to better learning, outperforms all previously reported dependency parsers for English and Chinese, yet is much faster.

Hard Constraints for Grammatical Function Labelling
Wolfgang Seeker, Ines Rehbein, Jonas Kuhn and Josef van Genabith
For languages with (semi-) free word order (such as German), labelling grammatical functions on top of phrase-structural constituent analyses is crucial for making them interpretable. Unfortunately, most statistical classifiers consider only local information for function labelling and fail to capture important restrictions on the distribution of core argument functions such as subject, object etc., namely that there is at most one subject (etc.) per clause. We augment a statistical classifier with an integer linear program imposing hard linguistic constraints on the solution space output by the classifier, capturing global distributional restrictions. We show that this improves labelling quality, in particular for argument grammatical functions, in an intrinsic evaluation, and, importantly, grammar coverage for treebank- based (Lexical-Functional) grammar acquisition and parsing, in an extrinsic evaluation.

Simple, Accurate Parsing with an All-Fragments Grammar
Mohit Bansal and Dan Klein
We present a simple but accurate parser which exploits both large tree fragments and symbol refinement. We parse with all fragments of the training set, in contrast to much recent work on tree selection in data-oriented parsing and tree-substitution grammar learning. We require only simple, deterministic grammar symbol refinement, in contrast to recent work on latent symbol refinement. Moreover, our parser requires no explicit lexicon machinery, instead parsing input sentences as character streams. Despite its simplicity, our parser achieves accuracies of over 88% F1 on the standard English WSJ task, which is competitive with substantially more complicated state-of-the-art lexicalized and latent-variable parsers. Additional specific contributions center on making implicit all-fragments parsing efficient, including a coarse-to-fine inference scheme and a new graph encoding.

Joint Syntactic and Semantic Parsing of Chinese
Junhui Li, Guodong Zhou and Hwee Tou Ng
This paper explores joint syntactic and semantic parsing of Chinese to further improve the performance of both syntactic and semantic parsing, in particular the performance of semantic parsing (in this paper, semantic role labeling). This is done from two levels. Firstly, an integrated parsing approach is proposed to integrate semantic parsing into the syntactic parsing process. Secondly, semantic information generated by semantic parsing is incorporated into the syntactic parsing model to better capture semantic information in syntactic parsing. Evaluation on Chinese TreeBank, Chinese PropBank, and Chinese NomBank shows that our integrated parsing approach outperforms the pipeline parsing approach on n-best parse trees, a natural extension of the widely

used pipeline parsing approach on the top-best parse tree. Moreover, it shows that incorporating semantic role-related information into the syntactic parsing model significantly improves the performance of both syntactic parsing and semantic parsing. To our best knowledge, this is the first research on exploring syntactic parsing and semantic role labeling for both verbal and nominal predicates in an integrated way.

Text Classification and Topic Models, 10:30–12:10, Venue A, Hall X

Cross-Language Text Classification Using Structural Correspondence Learning
Peter Prettenhofer and Benno Stein
We present a new approach to cross-language text classification that builds on structural correspondence learning, a recently proposed theory for domain adaptation. The approach uses unlabeled documents, along with a simple word translation oracle, in order to induce task-specific, cross-lingual word correspondences. We report on analyses that reveal quantitative insights about the use of unlabeled data and the complexity of inter-language correspondence modeling. We conduct experiments in the field of cross-language sentiment classification, employing English as source language, and German, French, and Japanese as target languages. The results are convincing; they demonstrate both the robustness and the competitiveness of the presented ideas.

Cross-Lingual Latent Topic Extraction
Duo Zhang, Qiaozhu Mei and ChengXiang Zhai
Probabilistic latent topic models have recently enjoyed much success in extracting and analyzing latent topics in text in an unsupervised way. One common deficiency of existing topic models, though, is that they would not work well for extracting cross-lingual latent topics simply because words in different language generally do not co-occur with each other. In this paper, we propose a way to incorporate a bilingual dictionary into a probabilistic topic model so that we can apply topic models to extract shared latent topics in text data of different languages. Specifically, we propose a new topic model called Probabilistic Cross-Lingual Latent Semantic Analysis (PCLSA) which extends the Probabilistic Latent Semantic Analysis (PLSA) model by regularizing its likelihood function with soft constraints defined based on a bilingual dictionary. Both qualitative and quantitative experimental results show that the PCLSA model can effectively extract cross-lingual latent topics from multilingual text data.

Topic Models for Word Sense Disambiguation and Token-Based Idiom Detection
Linlin Li, Benjamin Roth and Caroline Sporleder
This paper presents a probabilistic model for sense disambiguation which chooses the best sense based on the conditional probability of sense paraphrases given a context. We use a topic model to decompose this conditional probability into two conditional probabilities with latent variables. We propose three different instantiations of the model for solving sense disambiguation problems with different degrees of resource availability. The proposed models are tested on three different tasks: coarse-grained word sense disambiguation, fine-grained word sense disambiguation, and detection of literal vs. nonliteral usages of potentially idiomatic expressions. In all three cases, we outperform state-of-the-art systems either quantitatively or statistically significantly.

PCFGs, Topic Models, Adaptor Grammars and Learning Topical Collocations and the Structure of Proper Names
Mark Johnson
This paper establishes a connection between two apparently very different kinds of probabilistic models. Latent Dirichlet Allocation (LDA) models are used as "topic models" to produce a low-dimensional representation of documents, while Probabilistic Context-Free Grammars (PCFGs) define distributions over trees. The paper begins by showing that LDA topic models can be viewed as a special kind of PCFG, so Bayesian inference for PCFGs can be used to infer Topic Models as

well. Adaptor Grammars (AGs) are a hierarchical, non-parameteric Bayesian extension of PCFGs. Exploiting the close relationship between LDA and PCFGs just described, we propose two novel probabilistic models that combine insights from LDA and AG models. The first replaces the unigram component of LDA topic models with multi-word sequences or collocations generated by an AG. The second extension builds on the first one to learn aspects of the internal structure of proper names.

Psycholinguistics, 10:30–12:10, Venue A, Hall IX

A Cognitive Cost Model of Annotations Based on Eye-Tracking Data
Katrin Tomanek, Udo Hahn, Steffen Lohmann and Jürgen Ziegler
We report on an experiment to track complex decision points in linguistic meta-data annotation where the decision behavior of annotators is observed with an eye-tracking device. As experimental conditions we investigate different forms of textual context and linguistic complexity classes relative to syntax and semantics. Our data renders evidence that annotation performance depends on the semantic and syntactic complexity of the decision points and, more interestingly, indicates that full-scale context is mostly negligible – with the exception of semantic high-complexity cases. We then induce from this observational data a cognitively grounded cost model of linguistic meta-data annotations and compare it with existing non-cognitive models. Our data reveals that the cognitively founded model explains annotation costs (expressed in annotation time) more adequately than non-cognitive ones.

A Rational Model of Eye Movement Control in Reading
Klinton Bicknell and Roger Levy
A number of results in the study of real-time sentence comprehension have been explained by computational models as resulting from the rational use of probabilistic linguistic information. Many times, these hypotheses have been tested in reading by linking predictions about relative word difficulty to word-aggregated eye tracking measures such as go-past time. In this paper, we extend these results by asking to what extent reading is well-modeled as rational behavior at a finer level of analysis, predicting not aggregate measures, but the duration and location of each fixation. We present a new rational model of eye movement control in reading, the central assumption of which is that eye movement decisions are made to obtain noisy visual information as the reader performs Bayesian inference on the identities of the words in the sentence. As a case study, we present two simulations demonstrating that the model gives a rational explanation for between-word regressions.

The Influence of Discourse on Syntax: A Psycholinguistic Model of Sentence Processing
Amit Dubey
Probabilistic models of sentence comprehension are increasingly relevant to questions concerning human language processing. However, such models are often limited to syntactic factors. This paper introduces a novel sentence processing model that consists of a parser augmented with a probabilistic logic-based model of coreference resolution, which allows us to simulate how context interacts with syntax in a reading task. Our simulations show that a Weakly Interactive cognitive architecture can explain data which had been provided as evidence for the Strongly Interactive hypothesis.

Complexity Metrics in an Incremental Right-Corner Parser
Stephen Wu, Asaf Bachrach, Carlos Cardenas and William Schuler
Hierarchical HMM (HHMM) parsers make promising cognitive models: while they use a bounded model of working memory and pursue incremental hypotheses in parallel, they still achieve parsing accuracies competitive with chart-based techniques. This paper aims to validate that a right-corner HHMM parser is also able to produce complexity metrics, which quantify a reader's incre-

mental difficulty in understanding a sentence. Besides defining standard metrics in the HHMM framework, a new metric, embedding difference, is also proposed, which tests the hypothesis that HHMM store elements represents syntactic working memory. Results show that HHMM surprisal outperforms all other evaluated metrics in predicting reading times, and that embedding difference makes a significant, independent contribution.

Semantics 5, 10:30–12:10, Venue B, Lecture Hall 3

"Ask Not What Textual Entailment Can Do for You..."
Mark Sammons, V.G.Vinod Vydiswaran and Dan Roth
We challenge the NLP community to participate in a large-scale, distributed effort to design and build resources for developing and evaluating solutions to new and existing NLP tasks in the context of Recognizing Textual Entailment. We argue that the single global label with which RTE examples are annotated is insufficient to effectively evaluate RTE system performance; to promote research on smaller, related NLP tasks, we believe more detailed annotation and evaluation are needed, and that this effort will benefit not just RTE researchers, but the NLP community as a whole. We use insights from successful RTE systems to propose a model for identifying and annotating textual inference phenomena in textual entailment examples, and we present the results of a pilot annotation study that show this model is feasible and the results immediately useful.

Assessing the Role of Discourse References in Entailment Inference
Shachar Mirkin, Ido Dagan and Sebastian Pado
Discourse references, notably coreference and bridging, play an important role in many text understanding applications, but their impact on textual entailment is yet to be systematically understood. On the basis of an in-depth analysis of entailment instances, we argue that discourse references have the potential of substantially improving textual entailment recognition, and identify a number of research directions towards this goal.

Global Learning of Focused Entailment Graphs
Jonathan Berant, Ido Dagan and Jacob Goldberger
We propose a global algorithm for learning entailment relations between predicates. We define a graph structure over predicates that represents entailment relations as directed edges, and use a global transitivity constraint on the graph to learn the optimal set of edges, by formulating the optimization problem as an Integer Linear Program. We motivate this graph with an application that provides a hierarchical summary for a set of propositions that focus on a target concept, and show that our global algorithm improves performance by more than 10% over baseline algorithms.

Modeling Semantic Relevance for Question-Answer Pairs in Web Social Communities
Baoxun Wang, Xiaolong Wang, Chengjie Sun, Bingquan Liu and Lin Sun
Quantifying the semantic relevance between questions and their candidate answers is essential to answer detection in social media corpora. In this paper, a deep belief network is proposed to model the semantic relevance for question-answer pairs. Observing the textual similarity between the community-driven question-answering (cQA) dataset and the forum dataset, we present a novel learning strategy to promote the performance of our method on the social community datasets without hand-annotating work. The experimental results show that our method outperforms the traditional approaches on both the cQA and the forum corpora.

Multimodal, 10:30–12:10, Venue B, Lecture Hall 4

How Many Words Is a Picture Worth? Automatic Caption Generation for News Images
Yansong Feng and Mirella Lapata
In this paper we tackle the problem of automatic caption generation for news images. Our ap-

proach leverages the vast resource of pictures available on the web and the fact that many of them are captioned. Inspired by recent work in summarization, we propose extractive and abstractive caption generation models. They both operate over the output of a probabilistic image annotation model that preprocesses the pictures and suggests keywords to describe their content. Experimental results show that an abstractive model defined over phrases is superior to extractive methods.

Generating Image Descriptions Using Dependency Relational Patterns
Ahmet Aker and Robert Gaizauskas
This paper presents a novel approach to automatic captioning of geo-tagged images by summarizing multiple web-documents that contain information related to an image's location. The summarizer is biased by dependency pattern models towards sentences which contain features typically provided for different scene types such as those of churches, bridges, etc. Our results show that summaries biased by dependency pattern models lead to significantly higher ROUGE scores than both n-gram language models reported in previous work and also Wikipedia baseline summaries. Summaries generated using dependency patterns also lead to more readable summaries than those generated without dependency patterns.

Incorporating Extra-Linguistic Information into Reference Resolution in Collaborative Task Dialogue
Ryu Iida, Syumpei Kobayashi and Takenobu Tokunaga
This paper proposes an approach to reference resolution in situated dialogues by exploiting extra-linguistic information. Recently, investigations of referential behaviours involved in situations in the real world have received increasing attention by researchers (Di Eugenio et al., 2000; Byron, 2005; van Deemter, 2007; Spanger et al., 2009). In order to create an accurate reference resolution model, we need to handle extra-linguistic information as well as textual information examined by existing approaches (Soon et al., 2001; Ng and Cardie, 2002, etc.). In this paper, we incorporate extra-linguistic information into an existing corpus-based reference resolution model, and investigate its effects on reference resolution problems within a corpus of Japanese dialogues. The results demonstrate that our proposed model achieves an accuracy of 79.0% for this task.

Reading between the Lines: Learning to Map High-Level Instructions to Commands
S.R.K. Branavan, Luke Zettlemoyer and Regina Barzilay
In this paper, we address the task of mapping high-level instructions to commands in an external environment. Processing these instructions is challenging—they posit goals to be achieved without specifying the steps required to complete them. We describe a method that fills in missing information using an automatically derived environment model that encodes states, transitions, and commands that cause these transitions to happen. We present an efficient approximate approach for learning this environment model as part of a policy-gradient reinforcement learning algorithm for text interpretation. This design enables learning for mapping high-level instructions, which previous statistical methods cannot handle.

Unsupervised Parsing and Grammar Induction, 14:30–15:45, Venue A, Aula

Profiting from Mark-Up: Hyper-Text Annotations for Guided Parsing
Valentin I. Spitkovsky, Daniel Jurafsky and Hiyan Alshawi
We show how web mark-up can be used to improve unsupervised dependency parsing. Starting from raw bracketings of four common HTML tags (anchors, bold, italics and underlines), we refine approximate partial phrase boundaries to yield accurate parsing constraints. Conversion procedures fall out of our linguistic analysis of a newly available million-word hyper-text corpus. We demonstrate that derived constraints aid grammar induction by training Klein and Manning's

Dependency Model with Valence (DMV) on this data set: parsing accuracy on Section 23 (all sentences) of the Wall Street Journal corpus jumps to 50.4%, beating previous state-of-the-art by more than 5%. Web-scale experiments show that the DMV, perhaps because it is unlexicalized, does not benefit from orders of magnitude more annotated but noisier data. Our model, trained on a single blog, generalizes to 53.3% accuracy out-of-domain, against the Brown corpus — nearly 10% higher than the previous published best. The fact that web mark-up strongly correlates with syntactic structure may have broad applicability in NLP.

Phylogenetic Grammar Induction
Taylor Berg-Kirkpatrick and Dan Klein
We present an approach to multilingual grammar induction that exploits a phylogeny-structured model of parameter drift. Our method does not require any translated texts or token-level alignments. Instead, the phylogenetic prior couples languages at a parameter level. Joint induction in the multilingual model substantially outperforms independent learning, with larger gains both from more articulated phylogenies and as well as from increasing numbers of languages. Across eight languages, the multilingual approach gives error reductions over the standard monolingual DMV averaging 21.1% and reaching as high as 39%.

Improved Unsupervised POS Induction through Prototype Discovery
Omri Abend, Roi Reichart and Ari Rappoport
We present a novel fully unsupervised algorithm for POS induction from plain text, motivated by the cognitive notion of prototypes. The algorithm first identifies landmark clusters of words, serving as the cores of the induced POS categories. The rest of the words are subsequently mapped to these clusters. We utilize morphological and distributional representations computed in a fully unsupervised manner. We evaluate our algorithm on English and German, achieving the best reported results for this task.

Information Extraction 3, 14:30–15:45, Venue A, Hall X

Extraction and Approximation of Numerical Attributes from the Web
Dmitry Davidov and Ari Rappoport
We present a novel framework for automated extraction and approximation of numerical object attributes such as height and weight from the Web. Given an object-attribute pair, we discover and analyze attribute information for a set of comparable objects in order to infer the desired value. This allows us to approximate the desired numerical values even when no exact values can be found in the text. Our framework makes use of relation defining patterns and WordNet similarity information. First, we obtain from the Web and WordNet a list of terms similar to the given object. Then we retrieve attribute values for each term in this list, and information that allows us to compare different objects in the list and to infer the attribute value range. Finally, we combine the retrieved data for all terms from the list to select or approximate the requested value. We evaluate our method using automated question answering, WordNet enrichment, and comparison with answers given in Wikipedia and by leading search engines. In all of these, our framework provides a significant improvement.

Learning Word-Class Lattices for Definition and Hypernym Extraction
Roberto Navigli and Paola Velardi
Definition extraction is the task of automatically identifying definitional sentences within texts. The task has proven useful in many research areas including ontology learning, relation extraction and question answering. However, current approaches, mostly focused on lexico-syntactic patterns, suffer from both low recall and precision, as definitional sentences occur in highly variable syntactic structures. In this paper, we propose Word-Class Lattices (WCLs), a generalization of

word lattices that we use to model textual definitions. Lattices are learned from a large dataset of definitions from Wikipedia. Our method is applied to the task of definition and hypernym extraction and compares favorably to other pattern generalization methods proposed in the literature.

On Learning Subtypes of the Part-Whole Relation: Do Not Mix Your Seeds
Ashwin Ittoo and Gosse Bouma
An important relation in information extraction is the part-whole relation. Ontological studies mention several types of this relation. In this paper, we show that the traditional practice of initializing minimally-supervised algorithms with a single set that mixes seeds of different types fails to capture the wide variety of part-whole patterns and tuples. The results obtained with mixed seeds ultimately converge to one of the part-whole relation types. We also demonstrate that all the different types of part-whole relations can still be discovered, regardless of the type characterized by the initializing seeds. We performed our experiments with a state-of the art information extraction algorithm.

Information Retrieval, 14:30–15:45, Venue A, Hall IX

Understanding the Semantic Structure of Noun Phrase Queries
Xiao Li
Determining the semantic intent of web queries not only involves identifying their semantic class, which is a primary focus of previous works, but also understanding their semantic structure. In this work, we formally define the semantic structure of noun phrase queries as comprised of intent heads and intent modifiers. We present methods that automatically identify these constituents as well as their semantic roles based on Markov and semi-Markov conditional random fields. We show that the use of semantic features and syntactic features significantly contribute to improving the understanding performance.

Multilingual Pseudo-Relevance Feedback: Performance Study of Assisting Languages
Manoj Kumar Chinnakotla, Karthik Raman and Pushpak Bhattacharyya
In a previous work of ours Chinnakotla et al. (2010) we introduced a novel framework for Pseudo-Relevance Feedback (PRF) called MultiPRF. Given a query in one language called Source, we used English as the Assisting Language to improve the performance of PRF for the source language. MulitiPRF showed remarkable improvement over plain Model Based Feedback (MBF) uniformly for 4 languages, viz., French, German, Hungarian and Finnish with English as the assisting language. This fact inspired us to study the effect of any source-assistant pair on MultiPRF performance from out of a set of languages with widely different characteristics, viz., Dutch, English, Finnish, French, German and Spanish. Carrying this further, we looked into the effect of using two assisting languages together on PRF. The present paper is a report of these investigations, their results and conclusions drawn therefrom. While performance improvement on MultiPRF is observed whatever the assisting language and whatever the source, observations are mixed when two assisting languages are used simultaneously. Interestingly, the performance improvement is more pronounced when the source and assisting languages are closely related, e.g., French and Spanish.

Wikipedia as Sense Inventory to Improve Diversity in Web Search Results
Celina Santamaría, Julio Gonzalo and Javier Artiles
Is it possible to use sense inventories to improve Web search results diversity for one word queries? To answer this question, we focus on two broad-coverage lexical resources of a different nature: WordNet, as a de-facto standard used in Word Sense Disambiguation experiments; and Wikipedia, as a large coverage, updated encyclopaedic resource which may have a better coverage of relevant senses in Web pages. Our results indicate that (i) Wikipedia has a much better coverage of search results, (ii) the distribution of senses in search results can be estimated using the internal graph

structure of the Wikipedia and the relative number of visits received by each sense in Wikipedia, and (iii) associating Web pages to Wikipedia senses with simple and efficient algorithms, we can produce modified rankings that cover 70% more Wikipedia senses than the original search engine rankings.

Sentiment 3, 14:30–15:45, Venue B, Lecture Hall 3

A Unified Graph Model for Sentence-Based Opinion Retrieval
Binyang Li, Lanjun Zhou, Shi Feng and Kam-Fai Wong
There is a growing research interest in opinion retrieval as on-line users' opinions are becom-ing more and more popular in business, social network, etc. Practically speaking, the goal of opinion retrieval is to retrieve documents, which entail opinions or comments, relevant to a target specified by the user's query. A fun-damental challenge in opinion retrieval is in-formation representation. Existing research fo-cuses on document-based approaches and documents are represented by bag-of-word. However, due to loss of contextual information, this representation fails to capture the associative information between an opinion and its corresponding target and it cannot distinguish different degrees of a sentiment word when associated with different targets, which in turn seriously affects opinion retrieval performance. In this paper, we propose a sentence-based approach and define a new information representation, topic-sentiment word pair, to capture intra-sentence contextual information between the opinion and its target. Additionally, we consider inter-sentence information to capture the relationships among the opinions on the same topic. Finally, two types of information are combined in a novel unified graph-based model, which can effectively rank the docu-ments. Compared with existing approaches, experimental results on the COAE08 dataset show that our graph-based model has achieved significant improvement.

Generating Fine-Grained Reviews of Songs from Album Reviews
Swati Tata and Barbara Di Eugenio
Music Recommendation Systems often recommend individual songs, as opposed to entire albums. The challenge is to generate reviews for each song, since only full album reviews are available online. We developed a summarizer that combines information extraction and generation techniques to produce summaries of reviews of individual songs. We present an intrinsic evaluation of the extraction components, and of the informativeness of the summaries; and a user study of the impact of the song review summaries on users' decision making processes. Users were able to make quicker and more informed decisions when presented with the summary as compared to the full album review.

A Study of Information Retrieval Weighting Schemes for Sentiment Analysis
Georgios Paltoglou and Mike Thelwall
Most sentiment analysis approaches use as baseline a support vector machines (SVM) classifier with binary unigram weights. In this paper, we explore whether more sophisticated feature weighting schemes from Information Retrieval can enhance classification accuracy. We show that variants of the classic tf.idf scheme adapted to sentiment analysis provide significant increases in accuracy, especially when using a sublinear function for term frequency weights and document frequency smoothing. The techniques are tested on a wide selection of data sets and produce the best accuracy to our knowledge.

Discourse 2, 14:30–15:45, Venue B, Lecture Hall 4

Supervised Noun Phrase Coreference Research: The First Fifteen Years
Vincent Ng
The research focus of computational coreference resolution has exhibited a shift from heuristic approaches to machine learning approaches in the past decade. This paper surveys the major

milestones in supervised coreference research since its inception fifteen years ago.

Unsupervised Event Coreference Resolution with Rich Linguistic Features
Cosmin Bejan and Sanda Harabagiu
This paper examines how a new class of nonparametric Bayesian models can be effectively applied to an open-domain event coreference task. Designed with the purpose of clustering complex linguistic objects, these models consider a potentially infinite number of features and categorical outcomes. The evaluation performed for solving both within- and cross-document event coreference shows significant improvements of the models when compared against two baselines for this task.

Coreference Resolution across Corpora: Languages, Coding Schemes, and Preprocessing Information
Marta Recasens and Eduard Hovy
This paper explores the effect that different corpus configurations have on the performance of a coreference resolution system, as measured by MUC, B-CUBED, and CEAF. By varying separately three parameters (language, annotation scheme, and preprocessing information) and applying the same coreference resolution system, the strong bonds between system and corpus are demonstrated. The experiments reveal problems in coreference resolution evaluation relating to task definition, coding schemes, and features. They also expose systematic biases in the coreference evaluation metrics. We show that system comparison is only possible when corpus parameters are in exact agreement.

Translation 5, 16:15–17:30, Venue A, Aula

Constituency to Dependency Translation with Forests
Haitao Mi and Qun Liu
Tree-to-string systems (and their forest-based extensions) have gained steady popularity thanks to their simplicity and efficiency, but there is a major limitation: they are unable to guarantee the grammaticality of output, which is explicitly modeled in string-to-tree systems via target-side syntax. We thus propose to combine the advantages of both, and present a novel constituency-to-dependency transla- tion model, which uses constituency forests on the source side to direct the translation, and dependency trees on the target side (as a language model) to ensure grammaticality. Medium-scale experiments show an abso- lute and statistically significant improvement of +0.7 BLEU points over a state-of-the-art forest-based tree-to-string system even with fewer rules. This is also the first time that a tree-to-tree model can surpass tree-to-string counterparts.

Learning to Translate with Source and Target Syntax
David Chiang
Statistical translation models that try to capture the recursive structure of language have been widely adopted over the last few years. These models make use of varying amounts of information from linguistic theory: some use none at all, some use information about the grammar of the target language, some use information about the grammar of the source language. But progress has been slower on translation models that are able to learn the relationship between the grammars of both the source and target language. We discuss the reasons why this has been a challenge, review existing attempts to meet this challenge, and show how some old and new ideas can be combined into a simple approach that uses both source and target syntax for significant improvements in translation accuracy.

Discriminative Modeling of Extraction Sets for Machine Translation
John DeNero and Dan Klein
We present a discriminative model that directly predicts which set of phrasal translation rules

should be extracted from a sentence pair. Our model scores extraction sets: nested collections of all the overlapping phrase pairs consistent with an underlying word alignment. Extraction set models provide two principle advantages over word-factored alignment models. First, we can incorporate features on phrase pairs, in addition to word links. Second, we can optimize for an extraction-based loss function that relates directly to the end task of generating translations. Our model gives improvements in alignment quality relative to state-of-the-art unsupervised and supervised baselines, as well as providing up to a 1.4 improvement in BLEU score in Chinese-to-English translation experiments.

Information Extraction 4, 16:15–17:30, Venue A, Hall X

Detecting Experiences from Weblogs
Keun Chan Park, Yoonjae Jeong and Sung Hyon Myaeng
Weblogs are a source of human activity knowledge comprising valuable information such as facts, opinions and personal experiences. In this paper, we propose a method for mining personal experiences from a large set of weblogs. We define experience as knowledge embedded in a collection of activities or events which an individual or group has actually undergone. Based on an observation that experience-revealing sentences have a certain linguistic style, we formulate the problem of detecting experience as a classification task using various linguistic features. We also present an activity lexicon construction method based on theories of lexical semantics. Our results demonstrate that the activity lexicon plays a pivotal role among selected features in the classification performance and shows that our proposed method outperforms the baseline significantly.

Experiments in Graph-Based Semi-Supervised Learning Methods for Class-Instance Acquisition
Partha Pratim Talukdar and Fernando Pereira
Graph-based semi-supervised learning (SSL) algorithms have been successfully used to extract class-instance pairs from large unstructured and structured text collections. However, a careful comparison of different graph-based SSL algorithms on that task has been lacking. We compare three graph-based SSL algorithms for class-instance acquisition on a variety of graphs constructed from different domains. We find that the recently proposed MAD algorithm is the most effective. We also show that class-instance extraction can be significantly improved by adding semantic information in the form of instance-attribute edges derived from an independently developed knowledge base. All of our code and data will be made publicly available to encourage reproducible research in this area.

Learning Arguments and Supertypes of Semantic Relations Using Recursive Patterns
Zornitsa Kozareva and Eduard Hovy
A challenging problem in open information extraction and text mining is the learning of the selectional restrictions of semantic relations. We propose a minimally supervised bootstrapping algorithm that uses a single seed and a recursive lexico-syntactic pattern to learn the arguments and the supertypes of a diverse set of semantic relations from the Web. We evaluate the performance of our algorithm on multiple semantic relations that can be expressed using "verb", "noun", and "verb prep" lexico-syntactic patterns. Human-based evaluation shows that the accuracy of the harvested information is about 90%. We also compare our results with existing knowledge base to outline the similarities and differences of the granularity and diversity of the harvested knowledge.

Parsing and Grammars, 16:15–17:30, Venue A, Hall IX

A Transition-Based Parser for 2-Planar Dependency Structures
Carlos Gómez-Rodríguez and Joakim Nivre

Finding a class of structures that is rich enough for adequate linguistic representation yet restricted enough for efficient computational processing is an important problem for dependency parsing. In this paper, we present a transition system for 2-planar dependency trees – trees that can be decomposed into at most two planar graphs – and show that it can be used to implement a classifier-based parser that runs in linear time and outperforms a state-of-the-art transition-based parser on four data sets from the CoNLL-X shared task. In addition, we present an efficient method for determining whether an arbitrary tree is 2-planar and show that 99% or more of the trees in existing treebanks are 2-planar.

Viterbi Training for PCFGs: Hardness Results and Competitiveness of Uniform Initialization
Shay Cohen and Noah A. Smith
We consider the search for a maximum likelihood assignment of hidden derivations and grammar weights for a probabilistic context-free grammar, the problem approximately solved by "Viterbi training." We show that solving and even approximating Viterbi training for PCFGs is NP-hard. We motivate the use of uniform-at-random initialization for Viterbi EM as an optimal initializer in absence of further information about the correct model parameters, providing an approximate bound on the log-likelihood.

A Generalized-Zero-Preserving Method for Compact Encoding of Concept Lattices
Matthew Skala, Victoria Krakovna, János Kramár and Gerald Penn
Constructing an encoding of a concept lattice using short bit vectors allows for efficient computation of join operations on the lattice. Join is the central operation any unification-based parser must support. We extend the traditional bit vector encoding, which represents join failure using the zero vector, to count any vector with less than a fixed number of one bits as failure. This allows non-joinable elements to share bits, resulting in a smaller vector size. A constraint solver is used to construct the encoding, and a variety of techniques are employed to find near-optimal solutions and handle timeouts. An evaluation is provided comparing the extended representation of failure with traditional bit vector techniques.

Word Sense Disambiguation, 16:15–17:30, Venue B, Lecture Hall 3

Knowledge-Rich Word Sense Disambiguation Rivaling Supervised Systems
Simone Paolo Ponzetto and Roberto Navigli
One of the main obstacles to high-performance Word Sense Disambiguation (WSD) is the knowledge acquisition bottleneck. In this paper, we present a methodology to automatically extend WordNet with large amounts of semantic relations from an encyclopedic resource, namely Wikipedia. We show that, when provided with a vast amount of high-quality semantic relations, simple knowledge-lean disambiguation algorithms compete with state-of-the-art supervised WSD systems in a coarse-grained all-words setting and outperform them on gold-standard domain-specific datasets.

All Words Domain Adapted WSD: Finding a Middle Ground between Supervision and Unsupervision
Mitesh Khapra, Anup Kulkarni, Saurabh Sohoney and Pushpak Bhattacharyya
In spite of decades of research on Word Sense Disambiguation (WSD), all-words general purpose WSD has remained a distant goal. Many supervised WSD systems have been built, but the effort of creating the training corpus – annotated sense marked corpora – has always been a matter of concern. Therefore, attempts have been made to develop unsupervised and knowledge based techniques for WSD which do not need sense marked corpora. However such approaches have not proved effective, since they typically do not better Wordnet first sense baseline accuracy. Our research reported here proposes to stick to the supervised approach, but with far less demand on annotation. We show that if we have ANY sense marked corpora, be it from mixed domain or a

specific domain, a small amount of annotation in ANY other domain can deliver the goods almost as if exhaustive sense marking were available in that domain. We have tested our approach across Tourism and Health domain corpora, using also the well known mixed domain SemCor corpus. Accuracy figures close to "self domain" training lend credence to the viability of our approach. Our contribution thus lies in finding a convenient middle ground between pure supervised and pure unsupervised WSD. Finally, our approach is not restricted to any specific set of target words, a departure from a commonly observed practice in domain specific WSD.

Combining Orthogonal Monolingual and Multilingual Sources of Evidence for All Words WSD
Weiwei Guo and Mona Diab
Word Sense Disambiguation remains one of the most complex problems facing computational linguists to date. In this paper we present a system that combines evidence from a monolingual WSD system together with that from a multilingual WSD system to yield state of the art performance on standard All-Words data sets. The monolingual system is based on a modification of the graph based state of the art algorithm In-Degree. The multilingual system is an improvement over an All-Words unsupervised approach, SALAAM. SALAAM exploits multilingual evidence as a means of disambiguation. In this paper, we present modifications to both of the original approaches and then their combination. We finally report the highest results obtained to date on the SENSEVAL 2 standard data set using an unsupervised method, we achieve an overall F measure of 64.58 using a voting scheme.

Generation, 16:15–17:30, Venue B, Lecture Hall 4

Phrase-Based Statistical Language Generation Using Graphical Models and Active Learning
Francois Mairesse, Milica Gasic, Filip Jurcicek, Simon Keizer, Blaise Thomson, Kai Yu and Steve Young
Most previous work on trainable language generation has focused on two paradigms: (a) using a statistical model to rank a set of generated utterances, or (b) using statistics to inform the generation decision process. Both approaches rely on the existence of a handcrafted generator, which limits their scalability to new domains. This paper presents BAGEL, a statistical language generator which uses dynamic Bayesian networks to learn from semantically-aligned data produced by untrained annotators. A human evaluation shows that BAGEL can generate natural and informative utterances from unseen inputs in the information presentation domain. Additionally, generation performance on sparse datasets is improved significantly by using certainty-based active learning, yielding ratings close to the human gold standard with a fraction of the data.

Plot Induction and Evolutionary Search for Story Generation
Neil McIntyre and Mirella Lapata
In this paper we develop a story generator that leverages knowledge inherent in corpora without requiring extensive manual involvement. A key feature in our approach is the reliance on a story planner which we acquire automatically by recording events, their participants, and their precedence relationships in a training corpus. Contrary to previous work our system does not follow a generate-and-rank architecture. Instead, we employ evolutionary search techniques to explore the space of possible stories which we argue are well suited to the story generation task. Experiments on generating simple children's stories show that our system outperforms previous data-driven approaches.

Automated Planning for Situated Natural Language Generation
Konstantina Garoufi and Alexander Koller
We present a natural language generation approach which models, exploits, and manipulates the non-linguistic context in situated communication, using techniques from AI planning. We show how to generate instructions which deliberately guide the hearer to a location that is convenient for the generation of simple referring expressions, and how to generate referring expressions with

context-dependent adjectives. We implement and evaluate our approach in the framework of the Challenge on Generating Instructions in Virtual Environments, finding that it performs well even under the constraints of real-time generation.

Best Paper Session, 17:40–18:15, Venue A, Aula

Beyond NomBank: A Study of Implicit Arguments for Nominal Predicates

Matthew Gerber and Joyce Chai

Despite its substantial coverage, NomBank does not account for all within-sentence arguments and ignores extra-sentential arguments altogether. These arguments, which we call implicit, are important to semantic processing, and their recovery could potentially benefit many NLP applications. We present a study of implicit arguments for a select group of frequent nominal predicates. We show that implicit arguments are pervasive for these predicates, adding 65% to the coverage of NomBank. We demonstrate the feasibility of recovering implicit arguments with a supervised classification model. Our results and analyses provide a baseline for future work on this emerging task.

13 CoNLL-2010 Abstracts

CoNLL-2010 Abstracts: Thursday, July 15

Session 1: Parsing

Improvements in Unsupervised Co-Occurrence-Based Parsing
Christian Hänig
This paper presents an algorithm for unsupervised co-occurrence based parsing that improves and extends existing approaches. The proposed algorithm induces a context-free grammar of the language in question in an iterative manner. The resulting structure of a sentence will be given as a hierarchical arrangement of constituents. Although this algorithm does not use any a priori knowledge about the language, it is able to detect heads, modifiers and a phrase type's different compound composition possibilities. For evaluation purposes, the algorithm is applied to manually annotated part-of-speech tags (POS tags) as well as to word classes induced by an unsupervised part-of-speech tagger.

Viterbi Training Improves Unsupervised Dependency Parsing
Valentin I. Spitkovsky, Hiyan Alshawi, Daniel Jurafsky and Christopher D. Manning
We show that Viterbi (or "hard") EM is well-suited to unsupervised grammar induction. It is more accurate than standard inside-outside re-estimation (classic EM), significantly faster, and simpler. Our experiments with Klein and Manning's Dependency Model with Valence (DMV) attain state-of-the-art performance — 44.8% accuracy on Section 23 (all sentences) of the Wall Street Journal corpus — without clever initialization; with a good initializer, Viterbi training improves to 47.9%. This generalizes to the Brown corpus, our held-out set, where accuracy reaches 50.8% — a 7.5% gain over previous best results. We find that classic EM learns better from short sentences but cannot cope with longer ones, where Viterbi thrives. However, we explain that both algorithms optimize the wrong objectives and prove that there are fundamental disconnects between the likelihoods of sentences, best parses, and true parses, beyond the well-established discrepancies between likelihood, accuracy and extrinsic performance.

Driving Semantic Parsing from the World's Response
James Clarke, Dan Goldwasser, Ming-Wei Chang and Dan Roth
Current approaches to semantic parsing, the task of converting text to a formal meaning representation, rely on annotated training data mapping sentences to logical forms. Providing this supervision is a major bottleneck in scaling semantic parsers. This paper presents a new learning paradigm aimed at alleviating the supervision burden. We develop two novel learning algorithms capable of predicting complex structures which only rely on a binary feedback signal based on the context of an external world. In addition we reformulate the semantic parsing problem to reduce the dependency of the model on syntactic patterns, thus allowing our parser to scale better using less supervision. Our results surprisingly show that without using any annotated meaning representations learning with a weak feedback signal is capable of producing a parser that is competitive with fully supervised parsers.

Session 2: Grammar Induction

Efficient, Correct, Unsupervised Learning for Context-Sensitive Languages
Alexander Clark
A central problem for NLP is grammar induction: the development of unsupervised learning algorithms for syntax. In this paper we present a lattice-theoretic representation for natural language syntax, called Distributional Lattice Grammars. These representations are objective or empiricist, based on a generalisation of distributional learning, and are capable of representing all regular languages, some but not all context-free languages and some non-context-free languages. We present a simple algorithm for learning these grammars together with a complete self-contained proof of the correctness and efficiency of the algorithm.

Identifying Patterns for Unsupervised Grammar Induction
Jesús Santamaría and Lourdes Araujo
This paper describes a new method for unsupervised grammar induction based on the automatic extraction of certain patterns in the texts. Our starting hypothesis is that there exist some classes of words that function as separators, marking the beginning or the end of new constituents. Among these separators we distinguish those which trigger new levels in the parse tree. If we are able to detect these separators we can follow a very simple procedure to identify the constituents of a sentence by taking the classes of words between separators. This paper is devoted to describe the process that we have followed to automatically identify the set of separators from a corpus only annotated with Part-of-Speech (POS) tags. The proposed approach has allowed us to improve the results of previous proposals when parsing sentences from the Wall Street Journal corpus.

Learning Better Monolingual Models with Unannotated Bilingual Text
David Burkett, Slav Petrov, John Blitzer and Dan Klein
This work shows how to improve state-of-the-art monolingual natural language processing models using unannotated bilingual text. We build a multiview learning objective that enforces agreement between monolingual and bilingual models. In our method the first, monolingual view consists of supervised predictors learned separately for each language. The second, bilingual view consists of log-linear predictors learned over both languages on bilingual text. Our training procedure estimates the parameters of the bilingual model using the output of the monolingual model, and we show how to combine the two models to account for dependence between views. For the task of named entity recognition, using bilingual predictors increases F1 by 16.1% absolute over a supervised monolingual model, and retraining on bilingual predictions increases *monolingual* model F1 by 14.6%. For syntactic parsing, our bilingual predictor increases F1 by 2.1% absolute, and retraining a monolingual model on its output gives an improvement of 2.0%.

Invited Talk

Clueless: Explorations in Unsupervised, Knowledge-Lean Extraction of Lexical-Semantic Information
Lillian Lee
I will discuss two current projects on automatically extracting certain types of lexical-semantic information in settings wherein we can rely neither on annotations nor existing knowledge resources to provide us with clues. The name of the game in such settings is to find and leverage auxiliary sources of information. Why is it that if you *know* I'll give a silly talk, it follows that you know I'll give a talk, whereas if you *doubt* I'll give a good talk, it doesn't follow that you doubt I'll give a talk? This pair of examples shows that the word "doubt" exhibits a special but prevalent kind of behavior known as downward entailingness — the licensing of reasoning from supersets to subsets, so to speak, but not vice versa. The first project I'll describe is to identify words that are

downward entailing, a task that promises to enhance the performance of systems that engage in textual inference, and one that is quite challenging since it is difficult to characterize these items as a class and no corpus with downward-entailingness annotations exists. We are able to surmount these challenges by utilizing some insights from the linguistics literature regarding the relationship between downward entailing operators and what are known as negative polarity items — words such as "ever" or the idiom "have a clue" that tend to occur only in negative contexts. A cross-linguistic analysis indicates some potentially interesting connections to findings in linguistic typology. That previous paragraph was quite a mouthful, wasn't it? Wouldn't it be nice if it were written in plain English that was easier to understand? The second project I'll talk about, which has the eventual aim to make it possible to automatically simplify text, aims to learn lexical-level simplifications, such as "work together" for "collaborate". (This represents a complement to prior work, which focused on syntactic transformations, such as passive to active voice.) We exploit edit histories in Simple English Wikipedia for this task. This isn't as simple (ahem) as it might at first seem because Simple English Wikipedia and the usual Wikipedia are far from a perfect parallel corpus and because many edits in Simple Wikipedia do not constitute simplifications. We consider both explicitly modeling different kinds of operations and various types of bootstrapping, including as clues the comments Wikipedians sometimes leave when they edit. Joint work with Cristian Danescu-Niculescu-Mizil, Bo Pang, and Mark Yatskar.

Shared Task Session 1: Overview and Oral Presentations

The CoNLL 2010 Shared Task: Learning to Detect Hedges and their Scope in Natural Language Text

Richárd Farkas, Veronika Vincze, György Móra, János Csirik and György Szarvas

The CoNLL 2010 Shared Task was dedicated to the detection of uncertainty cues and their linguistic scope in natural language texts. The motivation behind this task was that distinguishing factual and uncertain information in texts is of essential importance in information extraction. This paper provides a general overview of the *Learning to detect hedges and their scope in natural language texts Shared Task*, including the annotation protocols of the training and evaluation datasets, the exact task definitions, the evaluation metrics employed and the overall results. The paper concludes with an analysis of the prominent approaches and an overview of the systems submitted to the Shared Task.

A Cascade Method for Detecting Hedges and their Scope in Natural Language Text

Buzhou Tang, Xiaolong Wang, Xuan Wang, Bo Yuan and Shixi Fan

Detecting hedges and their scope in natural language text is very important for information inference. In this paper, we present a system based on a cascade method for the CoNLL-2010 shared task. The system composes of two components: one for detecting hedges and another one for detecting their scope. For detecting hedges, we build a cascade subsystem. Firstly, a conditional random field (CRF) model and a large margin-based model are trained respectively. Then, we train another CRF model using the result of the first phase. For detecting the scope of hedges, a CRF model is trained according to the result of the first subtask. The experiments show that our system achieves 86.36% F-measure on Biological corpus and 55.05% F-measure on Wikipedia corpus for hedge detection, and 49.95% F-measure on Biological corpus for hedge scope detection. Among them, 86.36% is the best result on Biological corpus for hedge detection.

Detecting Speculative Language using Syntactic Dependencies and Logistic Regression

Andreas Vlachos and Mark Craven

In this paper we describe our approach to the CoNLL 2010 shared task on detecting speculative language in biomedical text. We treat the detection of sentences containing uncertain information (Task1) as a token classification task since the existence or absence of cues determines the sentence label. We distinguish words that have speculative and non-speculative meaning by employing

syntactic features as a proxy for their semantic content. In order to identify the scope of each cue (Task2), we learn a classifier that predicts whether each token of a sentence belongs to the scope of a given cue. The features in the classifier are based on the syntactic dependency path between the cue and the token. In both tasks, we use a Bayesian logistic regression classifier incorporating a sparsity-enforcing Laplace prior. Overall, the performance achieved is 85.21% F-score and 44.11% F-score in Task1 and Task2, respectively.

A Hedgehop over a Max-margin Framework using Hedge Cues
Maria Georgescul
In this paper, we describe the experimental settings we adopted in the context of the 2010 CoNLL shared task for detecting sentences containing uncertainty. The classification results reported on are obtained using discriminative learning with features essentially incorporating lexical information. Hyper-parameters are tuned for each domain: using BioScope training data for the biomedical domain and Wikipedia training data for the Wikipedia test set. By allowing an efficient handling of combinations of large-scale input features, the discriminative approach we adopted showed highly competitive empirical results for hedge detection on the Wikipedia dataset: our system is ranked as the first with an F-score of 60.17%.

Detecting Hedge Cues and their Scopes with Average Perceptron
Feng Ji, Xipeng Qiu and Xuanjing Huang
In this paper, we proposed a hedge detection method with average perceptron, which was used in the closed challenge in CoNLL 2010 Shared Task. There are two subtasks: (1) detecting uncertain sentences and (2) identifying the in-sentence scopes of hedge cues. We use the unified learning algorithm for both subtasks since that the hedge score of sentence can be decomposed into scores of the words, especially the hedge words. On the biomedical corpus, our methods achieved F-measure with 77.86% in detecting in-domain uncertain sentences, 77.44% in recognizing hedge cues, and 19.27% in identifying the scopes.

Memory-based Resolution of In-sentence Scopes of Hedge Cues
Roser Morante, Vincent Van Asch and Walter Daelemans
In this paper we describe the machine learning systems that we submitted to the CoNLL-2010 Shared Task on Learning to Detect Hedges and Their Scope in Natural Language Text. Task 1 on detecting uncertain information was performed by an SVM-based system to process the Wikipedia data and by a memory-based system to process the biological data. Task 2 on resolving in-sentence scopes of hedge cues, was performed by a memory-based system that relies on information from syntactic dependencies. This system scored the highest F1 (57.32) of Task 2.

Resolving Speculation: MaxEnt Cue Classification and Dependency-Based Scope Rules
Erik Velldal, Lilja Øvrelid and Stephan Oepen
This paper describes a hybrid, two-level approach for resolving hedge cues, the problem of the CoNLL 2010 shared task. First, a maximum entropy classifier is applied to identify cue words, using both syntactic and surface-oriented features. Second, a set of manually crafted rules, operating on dependency representations and the output of the classifier, is applied to resolve the scope of the hedge cues within the sentence. For both Task 1 and Task 2, our system participates in the stricter category of 'closed' or 'in-domain' systems.

Combining Manual Rules and Supervised Learning for Hedge Cue and Scope Detection
Marek Rei and Ted Briscoe
Hedge cues were detected using a supervised Conditional Random Field (CRF) classifier exploiting features from the RASP parser. The CRF's predictions were filtered using known cues and unseen instances were removed, increasing precision while retaining recall. Rules for scope detection, based on the grammatical relations of the sentence and the part-of-speech tag of the cue,

were manually developed. However, another supervised CRF classifier was used to refine these predictions. As a final step, scopes were constructed from the classifier output using a small set of post-processing rules. Development of the system revealed a number of issues with the annotation scheme adopted by the organisers.

CoNLL-2010 Abstracts: Friday, July 16

Invited Talk

Bayesian Hidden Markov Models and Extensions
Zoubin Ghahramani
Hidden Markov models (HMMs) are one of the cornerstones of time-series modelling. I will review HMMs, motivations for Bayesian approaches to inference in them, and our work on variational Bayesian learning. I will then focus on recent nonparametric extensions to HMMs. Traditionally, HMMs have a known structure with a fixed number of states and are trained using maximum likelihood techniques. The infinite HMM (iHMM) allows a potentially unbounded number of hidden states, letting the model use as many states as it needs for the data. The recent development of 'Beam Sampling' — an efficient inference algorithm for iHMMs based on dynamic programming — makes it possible to apply iHMMs to large problems. I will show some applications of iHMMs to unsupervised POS tagging and experiments with parallel and distributed implementations. I will also describe a factorial generalisation of the iHMM which makes it possible to have an unbounded number of binary state variables, and can be thought of as a time-series generalisation of the Indian buffet process. I will conclude with thoughts on future directions in Bayesian modelling of sequential data.

Joint Poster Session: Main conference and shared task posters

Improved Unsupervised POS Induction Using Intrinsic Clustering Quality and a Zipfian Constraint
Roi Reichart, Raanan Fattal and Ari Rappoport
Modern unsupervised POS taggers usually apply an optimization procedure to a non-convex function, and tend to converge to local maxima that are sensitive to starting conditions. The quality of the tagging induced by such algorithms is thus highly variable, and researchers report average results over several random initializations. Consequently, applications are not guaranteed to use an induced tagging of the quality reported for the algorithm. In this paper we address this issue using an unsupervised test for intrinsic clustering quality. We run a base tagger with different random initializations, and select the best tagging using the quality test. As a base tagger, we modify a leading unsupervised POS tagger (Clark, 2003) to constrain the distributions of word types across clusters to be Zipfian, allowing us to utilize a perplexity-based quality test. We show that the correlation between our quality test and gold standard-based tagging quality measures is high. Our results are better in most evaluation measures than all results reported in the literature for this task, and are always better than the Clark average results.

Syntactic and Semantic Structure for Opinion Expression Detection
Richard Johansson and Alessandro Moschitti
We demonstrate that relational features derived from dependency-syntactic and semantic role structures are useful for the task of detecting opinionated expressions in natural-language text, significantly improving over conventional models based on sequence labeling with local features. These features allow us to model the way opinionated expressions interact in a sentence over arbitrary distances. While the relational features make the prediction task more computationally expensive, we show that it can be tackled effectively by using a reranker. We evaluate a number of machine learning approaches for the reranker, and the best model results in a 10-point absolute improvement in soft recall on the MPQA corpus, while decreasing precision only slightly.

Type Level Clustering Evaluation: New Measures and a POS Induction Case Study
Roi Reichart, Omri Abend and Ari Rappoport

Clustering is a central technique in NLP. Consequently, clustering evaluation is of great importance. Many clustering algorithms are evaluated by their success in tagging corpus tokens. In this paper we discuss type level evaluation, which reflects class membership only and is independent of the token statistics of a particular reference corpus. Type level evaluation casts light on the merits of algorithms, and for some applications is a more natural measure of the algorithm's quality. We propose new type level evaluation measures that, contrary to existing measures, are applicable when items are polysemous, the common case in NLP. We demonstrate the benefits of our measures using a detailed case study, POS induction. We experiment with seven leading algorithms, obtaining useful insights and showing that token and type level measures can weakly or even negatively correlate, which underscores the fact that these two approaches reveal different aspects of clustering quality.

Recession Segmentation: Simpler Online Word Segmentation Using Limited Resources
Constantine Lignos and Charles Yang
In this paper we present a cognitively plausible approach to word segmentation that segments in an online fashion using only local information and a lexicon of previously segmented words. Unlike popular statistical optimization techniques, the learner uses structural information of the input syllables rather than distributional cues to segment words. We develop a memory model for the learner that like a child learner does not recall previously hypothesized words perfectly. The learner attains an F-score of 86.69% in ideal conditions and 85.05% when word recall is unreliable and stress in the input is reduced. These results demonstrate the power that a simple learner can have when paired with appropriate structural constraints on its hypotheses.

Computing Optimal Alignments for the IBM-3 Translation Model
Thomas Schoenemann
Prior work on training the IBM-3 translation model is based on suboptimal methods for computing Viterbi alignments. In this paper, we present the first method guaranteed to produce globally optimal alignments. This not only results in improved alignments, it also gives us the opportunity to evaluate the quality of standard hillclimbing methods. Indeed, hillclimbing works reasonably well in practice but still fails to find the global optimum for between 2% and 12% of all sentence pairs and the probabilities can be several tens of orders of magnitude away from the Viterbi alignment. By reformulating the alignment problem as an Integer Linear Program, we can use standard machinery from global optimization theory to compute the solutions. We use the well-known branch-and-cut method, but also show how it can be customized to the specific problem discussed in this paper. In fact, a large number of alignments can be excluded from the start without losing global optimality.

Semi-Supervised Recognition of Sarcasm in Twitter and Amazon
Dmitry Davidov, Oren Tsur and Ari Rappoport
Sarcasm is a form of speech act in which the speakers convey their message in an implicit way. The inherently ambiguous nature of sarcasm sometimes makes it hard even for humans to decide whether an utterance is sarcastic or not. Recognition of sarcasm can benefit many sentiment analysis NLP applications, such as review summarization, dialogue systems and review ranking systems. In this paper we experiment with semi-supervised sarcasm identification on two very different data sets: a collection of 5.9 million tweets collected from Twitter, and a collection of 66000 product reviews from Amazon. Using the Mechanical Turk we created a gold standard sample in which each sentence was tagged by 3 annotators, obtaining F-scores of 0.78 on the product reviews dataset and 0.83 on the Twitter dataset. We discuss the differences between the datasets and how the algorithm uses them (e.g., for the Amazon dataset the algorithm makes use of structured information). We also discuss the utility of Twitter #sarcasm hashtags for the task.

Learning Probabilistic Synchronous CFGs for Phrase-based Translation

Markos Mylonakis and Khalil Sima'an
Probabilistic phrase-based synchronous grammars are now considered promising devices for statistical machine translation because they can express reordering phenomena between pairs of languages. Learning these hierarchical, probabilistic devices from parallel corpora constitutes a major challenge, because of multiple latent model variables as well as the risk of data overfitting. This paper presents an effective method for learning a family of particular interest to MT, binary Synchronous Context-Free Grammars with inverted/monotone orientation (a.k.a. Binary ITG). A second contribution concerns devising a lexicalized phrase reordering mechanism that has complimentary strengths to Chiang's model. The latter conditions reordering decisions on the surrounding lexical context of phrases, whereas our mechanism works with the lexical content of phrase pairs (akin to standard phrase-based systems). Surprisingly, our experiments on French-English data show that our learning method applied to far simpler models exhibits performance indistinguishable from the Hiero system.

A Semi-Supervised Batch-Mode Active Learning Strategy for Improved Statistical Machine Translation

Sankaranarayanan Ananthakrishnan, Rohit Prasad, David Stallard and Prem Natarajan
The availability of substantial, in-domain parallel corpora is critical for the development of high-performance statistical machine translation (SMT) systems. Such corpora, however, are expensive to produce due to the labor intensive nature of manual translation. We propose to alleviate this problem with a novel, semi-supervised, batch-mode active learning strategy that attempts to maximize in-domain coverage by selecting sentences, which represent a balance between domain match, translation difficulty, and batch diversity. Simulation experiments on an English-to-Pashto translation task show that the proposed strategy not only outperforms the random selection baseline, but also traditional active learning techniques based on dissimilarity to existing training data. Our approach achieves a relative improvement of 45.9% in BLEU over the seed baseline, while the closest competitor gained only 24.8% with the same number of selected sentences.

Improving Word Alignment by Semi-supervised Ensemble

Shujian Huang, Kangxi Li, Xinyu Dai and Jiajun Chen
Supervised learning has been recently used to improve the performance of word alignment. However, due to the limited amount of labeled data, the performance of "pure" supervised learning, which only used labeled data, is limited. As a result, many existing methods employ features learnt from a large amount of unlabeled data to assist the task. In this paper, we propose a semi-supervised ensemble method to better incorporate both labeled and unlabeled data during learning. Firstly, we employ an ensemble learning framework, which effectively uses alignment results from different unsupervised alignment models. We then propose to use a semi-supervised learning method, namely Tri-training, to train classifiers using both labeled and unlabeled data collaboratively and further improve the result of ensemble learning. Experimental results show that our methods can substantially improve the quality of word alignment. The final translation quality of a phrase-based translation system is slightly improved, as well.

A Comparative Study of Bayesian Models for Unsupervised Sentiment Detection

Chenghua Lin, Yulan He and Richard Everson
This paper presents a comparative study of three closely related Bayesian models for unsupervised sentiment detection, namely, the latent sentiment model (LSM), the joint sentiment-topic (JST) model, and the Reverse-JST model. Extensive experiments have been conducted on two corpora, the movie review dataset and the multi-domain sentiment dataset. It has been found that while all the three models achieve either better or comparable performance on these two corpora when compared to the existing unsupervised sentiment classification approaches, both JST and Reverse-JST are able to extract sentiment-oriented topics. In addition, Reverse-JST always performs worse than JST suggesting that the JST model is more appropriate for joint sentiment topic detection.

A Hybrid Approach to Emotional Sentence Polarity and Intensity Classification
Jorge Carrillo de Albornoz, Laura Plaza and Pablo Gervás
In this paper, the authors present a new approach to sentence level sentiment analysis. The aim is to determine whether a sentence expresses a positive, negative or neutral sentiment, as well as its intensity. The method performs WSD over the words in the sentence in order to work with concepts rather than terms, and makes use of the knowledge in an affective lexicon to label these concepts with emotional categories. It also deals with the effect of negations and quantifiers on polarity and intensity analysis. An extensive evaluation in two different domains is performed in order to determine how the method behaves in 2-classes (positive and negative), 3-classes (positive, negative and neutral) and 5-classes (strongly negative, weakly negative, neutral, weakly positive and strongly positive) classification tasks. The results obtained compare favorably with those achieved by other systems addressing similar evaluations.

Cross-Caption Coreference Resolution for Automatic Image Understanding
Micah Hodosh, Peter Young, Cyrus Rashtchian and Julia Hockenmaier
In order to "understand" an image, it is necessary to identify not only the depicted entities, but also their attributes, relations between them and the actions they participate in. This information cannot be conveyed by simple keyword annotations. We have collected a corpus of 8108 "action" images associated each with five simple sentences describing their content and created a simple ontology of entity categories that appear in these images. In order to obtain a consistent semantic representation of the image content from these sentences, we need to first identify multiple mentions of the same entities. We present a hierarchical Bayesian model for cross-caption coreference resolution. We also evaluate how well the ontological types of the entities can be recovered.

Improved Natural Language Learning via Variance-Regularization Support Vector Machines
Shane Bergsma, Dekang Lin and Dale Schuurmans
We present a simple technique for learning better SVMs using fewer training examples. Rather than using the standard SVM regularization, we regularize toward low weight-variance. Our new SVM objective remains a convex quadratic function of the weights, and is therefore computationally no harder to optimize than a standard SVM. Variance regularization is shown to enable dramatic improvements in the learning rates of SVMs on three lexical disambiguation tasks.

Hedge Detection using the RelHunter Approach
Eraldo Fernandes, Carlos Crestana and Ruy Milidiú
RelHunter is a Machine Learning based method for the extraction of structured information from text. Here, we apply RelHunter to the Hedge Detection task, proposed as the CoNLL 2010 Shared Task. RelHunter's key design idea is to model the target structures as a relation over entities. The method decomposes the original task into three subtasks: (i) Entity Identification; (ii) Candidate Relation Generation; and (iii) Relation Recognition. In the Hedge Detection task, we define three types of entities: cue chunk, start scope token and end scope token. Hence, the Entity Identification subtask is further decomposed into three token classification subtasks, one for each entity type. In the Candidate Relation Generation subtask, we apply a simple procedure to generate a ternary candidate relation. Each instance in this relation represents a hedge candidate composed by a cue chunk, a start scope token and an end scope token. For the Relation Recognition subtask, we use a binary classifier to discriminate between true and false candidates. The four classifiers are trained with the Entropy Guided Transformation Learning algorithm. When compared to the other hedge detection systems of the CoNLL shared task, our scheme shows a competitive performance. The F-score of our system is 54.05 on the evaluation corpus.

A High-Precision Approach to Detecting Hedges and Their Scopes
Halil Kilicoglu and Sabine Bergler
We extend our prior work on speculative sentence recognition and speculation scope detection

in biomedical text to the CoNLL'10 Shared Task on Hedge Detection. In our participation, we sought to assess the extensibility and portability of our prior work, which relies on linguistic categorization and weighting of hedging cues and on syntactic patterns in which these cues play a role. For Task 1a, we tuned our categorization and weighting scheme to recognize hedging in biological text. By accommodating a small number of vagueness quantifiers, we were able to extend our methodology to detecting vague sentences in Wikipedia articles. We exploited constituent parse trees in addition to syntactic dependency relations in resolving hedging scope. Our results are competitive with those of closed-domain trained systems and demonstrate that our high-precision oriented methodology is extensible and portable.

Exploiting Rich Features for Detecting Hedges and Their Scope
Xinxin Li, Jianping Shen, Xiang Gao and Xuan Wang
This paper describes our system about detecting hedges and their scope in natural language texts for our participation in CoNLL2010 shared tasks. We formalize these two tasks as sequence labeling problems, and implement them using conditional random fields (CRFs) model. In the first task, we use a greedy forward procedure to select features for the classifier. These features include part-of-speech tag, word form, lemma, chunk tag of tokens in the sentence. In the second task, our system exploits rich syntactic features about dependency structures and phrase structures, which achieves a better performance than only using the flat sequence features. Our system achieves the third score in biological data set for the first task, and achieves 0.5265 F1 score for the second task.

Uncertainty Detection as Approximate Max-Margin Sequence Labelling
Oscar Täckström, Sumithra Velupillai, Martin Hassel, Gunnar Eriksson, Hercules Dalianis and Jussi Karlgren
This paper reports experiments for the CoNLL-2010 Shared Task on Learning to detect hedges and their scope in natural language text. We have addressed the experimental tasks as supervised linear maximum margin prediction prob- lems. For sentence level hedge detection in the biological domain we use an L1-regularised binary support vector machine, while for sentence level weasel detection in the Wikipedia domain, we use an L2-regularised approach. We model the in-sentence uncertainty cue and scope detection task as an L2-regularised approximate maximum margin sequence labelling problem, using the BIO-encoding. In addition to surface level features, we use a variety of linguistic features based on a functional dependency analysis. A greedy forward selection strategy is used in exploring the large set of potential features. Our official results for Task 1 for the biological domain were 0.852 F-score, for the Wikipedia set 0.5538 F-score. For Task 2, our official results were 0.0215 for the entire task with a score of 0.6249 for cue detection. After resolving errors and final bugs, our final results are for Task 1, biological: 0.788, Wikipedia: 0.577; Task 2: 0.396 and 0.785 for cues.

Hedge Detection and Scope Finding by Sequence Labeling with Procedural Feature Selection
Shaodian Zhang, Hai Zhao, Guodong Zhou and Bao-liang Lu
This paper presents a system which adopts a standard sequence labeling technique for hedge detection and scope finding. For hedge detection, we formulate it as a hedge labeling problem, while for hedge scope finding, we use a two-step labeling strategy, one for hedge labeling and the other for scope finding. In particular, various kinds of syntactic dependencies are systemically exploited and effectively integrated using a large-scale normalized feature selection method. Evaluation on the CoNLL-2010 shared task shows that our system achieves stable and competitive results for all the closed tasks. Furthermore, post-deadline experiments show that the performance can be much further improved using a sufficient feature selection.

Learning to Detect Hedges and their Scope using CRF
Qi Zhao, Chengjie Sun, Bingquan Liu and Yong Cheng
This paper presents an approach for extracting the hedge cues and their scopes in BioScope corpus

using two CRF models for CoNLL 2010 shared task. In the first task, the HCDic feature is proposed to improve the system performances, getting better performance (84.1% in F-score) than the baseline. The HCDic feature is also helpful to make use of cross-domain resources. The comparison of our methods based on between BioScope and Wikipedia corpus is given, which shows that ours are good at hedge cues detection in BioScope corpus but short at the in Wikipedia corpus. To detect the scope of hedge cues, we make rules to post process the text. For future work, we will look forward to constructing regulations for the HCDic to improve our system.

Exploiting Multi-Features to Detect Hedges and Their Scope in Biomedical Texts
Huiwei Zhou, Xiaoyan Li, Degen Huang, Zezhong Li and Yuansheng Yang
In this paper, we present a machine learning approach that detects hedge cues and their scope in biomedical texts. Identifying hedged information in texts is a kind of semantic filtering of texts and it is important since it could extract speculative information from factual information. In order to deal with the semantic analysis problem, various evidential features are proposed and integrated through a Conditional Random Fields (CRFs) model. Hedge cues that appear in the training dataset are regarded as keywords and employed as an important feature in hedge cue identification system. For the scope finding, we construct a CRF-based system and a syntactic pattern-based system, and compare their performances. Experiments using test data from CoNLL-2010 shared task show that our proposed method is robust. F-score of the biological hedge detection task and scope finding task achieves 86.32% and 54.18% in in-domain evaluations respectively.

A Lucene and Maximum Entropy Model Based Hedge Detection System
Lin Chen and Barbara Di Eugenio
This paper describes the approach to hedge detection we developed, in order to participate in the shared task at CoNLL 2010. A supervised learning approach is employed in our implementation. Hedge cue annotations in the training data are used as the seed to build a reliable hedge cue set. Maximum Entropy(MaxEnt) model is used as the learning technique to determine uncertainty. By making use of Apache Lucene, we are able to do fuzzy string match to extract hedge cues, and to incorporate part-of-speech(POS) tags in hedge cues. Not only can our system determine the certainty of the sentence, but is also able to find all the contained hedges. Our system was ranked third on the Wikipedia dataset. In later experiments with different parameters, we further improved our results, with a 0.612 F-score on the Wikipedia dataset, and a 0.802 F-score on the biological dataset.

HedgeHunter: A System for Hedge Detection and Uncertainty Classification
David Clausen
With the dramatic growth of scientific publishing, Information Extraction (IE) systems are becoming an increasingly important tool for large scale data analysis. Hedge detection and uncertainty classification are important components of a high precision IE system. This paper describes a two part supervised system which classifies words as hedge or non-hedged and sentences as certain or uncertain in biomedical and Wikipedia data. In the first stage, our system trains a logistic regression classifier to detect hedges based on lexical and Part-of-Speech collocation features. In the second stage, we use the output of the hedge classifier to generate sentence level features based on the number of hedge cues, the identity of hedge cues, and a Bag-of-Words feature vector to train a logistic regression classifier for sentence level uncertainty. With the resulting classification, an IE system can then discard facts and relations extracted from these sentences or treat them as appropriately doubtful. We present results for in domain training and testing and cross domain training and testing based on a simple union of training sets.

Exploiting CCG Structures with Tree Kernels for Speculation Detection
Liliana Paola Mamani Sanchez, Baoli Li and Carl Vogel
Our CoNLL-2010 speculative sentence detector disambiguates putative keywords based on the

following considerations: a speculative keyword may be composed of one or more word tokens; a speculative sentence may have one or more speculative keywords; and if a sentence contains at least one real speculative keyword, it is deemed speculative. A tree kernel classifier is used to assess whether a potential speculative keyword conveys speculation. We exploit information implicit in tree structures. For prediction efficiency, only a segment of the whole tree around a speculation keyword is considered, along with morphological features inside the segment and information about the containing document. A maximum entropy classifier is used for sentences not covered by the tree kernel classifier. Experiments on the Wikipedia data set show that our system achieves 0.55 F-measure (in-domain).

Uncertainty Learning using SVMs and CRFs
Vinodkumar Prabhakaran
In this work, I explore the use of SVMs and CRFs in the problem of predicting certainty in sentences. I consider this as a task of tagging uncertainty cues in context, for which I used lexical, wordlist-based and deep-syntactic features. Results show that the syntactic context of the tokens in conjunction with the wordlist-based features turned out to be useful in predicting uncertainty cues.

Features for Detecting Hedge Cues
Nobuyuki Shimizu and Hiroshi Nakagawa
We present a sequential labeling approach to hedge cue detection submitted to the CoNLL-2010 shared task, biological por- tion of task 1. Our main approach is as fol- lows. We make use of partial syntactic in- formation together with features obtained from the unlabeled corpus, and convert the t ask into a sequential BIO-tagging. If a cue is found, a sentence is classified as uncertain and certain otherwise. To ex- amine a large number of feature combi- nations, we employ a genetic algorithm. While some obtained features are difficult to interpret, they were shown to improve the performance of the final system.

A Simple Ensemble Method for Hedge Identification
Ferenc Szidarovszky, Illés Solt and Domonkos Tikk
We present in this paper a simple hedge identification method and its application on biomedical text. The problem at hand is a subtask of CoNLL 2010 shared task. Our solution consists of two classifiers, a statistical one and a CRF model, and a simple combination schema that combines their predictions. We report in detail on each component of our system and discuss the results. We also show that a more sophisticated combination schema could improve the F-score significantly.

A Baseline Approach for Detecting Sentences Containing Uncertainty
Erik Tjong Kim Sang
We apply a baseline approach to the CoNLL-2010 shared task data sets on hedge detection. Weights have been assigned to cue words marked in the training data based on their occurrences in certain and uncertain sentences. New sentences received scores that correspond with those of their best scoring cue word, if present. The best acceptance scores for uncertain sentences were determined using 10-fold cross validation on the training data. This approach performed reasonably on the shared task's biological (F=82.0) and Wikipedia (F=62.8) data sets.

Hedge Classification with Syntactic Dependency Features based on an Ensemble Classifier
Yi Zheng, Qifeng Dai, Qiming Luo and Enhong Chen
We present our CoNLL-2010 Shared Task system in the paper. The system operates in three steps: sequence labeling, syntactic de-pendency parsing, and classification. We have participated in the Shared Task 1. Our experi-mental results measured by the in-domain and cross-domain F-scores on the biological do-main are 81.11% and 67.99%, and on the Wikipedia domain 55.48% and 55.41%.

Session 3: Semantics and Information Extraction

Online Entropy-based Model of Lexical Category Acquisition
Grzegorz Chrupała and Afra Alishahi
Children learn a robust representation of lexical categories at a young age. We propose an incremental model of this process which efficiently groups words into lexical categories based on their local context using an information-theoretic criterion. We train our model on a corpus of child-directed speech from CHILDES and show that the model learns a fine-grained set of intuitive word categories. Furthermore, we propose a novel evaluation approach by comparing the efficiency of our induced categories against other category sets (including traditional part of speech tags) in a variety of language tasks. We show the categories induced by our model typically outperform the other category sets.

Tagging and Linking Web Forum Posts
Su Nam Kim, Li Wang and Timothy Baldwin
We propose a method for annotating post-to-post discourse structure in online user forum data, in the hopes of improving troubleshooting-oriented information access. We introduce the tasks of: (1) post classification, based on a novel dialogue act tag set; and (2) link classification. We also introduce three feature sets (structural features, post context features and semantic features) and experiment with three discriminative learners (maximum entropy, SVM-HMM and CRF). We achieve above-baseline results for both dialogue act and link classification, with interesting divergences in which feature sets perform well over the two sub-tasks, and go on to perform preliminary investigation of the interaction between post tagging and linking.

Joint Entity and Relation Extraction using Card-Pyramid Parsing
Rohit Kate and Raymond Mooney
Both entity and relation extraction can benefit from being performed jointly, allowing each task to correct the errors of the other. We present a new method for joint entity and relation extraction using a graph we call a "card-pyramid". This graph compactly encodes all possible entities and relations in a sentence, reducing the task of their joint extraction to jointly labeling its nodes. We give an efficient labeling algorithm that is analogous to parsing using dynamic programming. Experimental results show improved results for our joint extraction method compared to a pipelined approach.

Session 4: Machine learning

Distributed Asynchronous Online Learning for Natural Language Processing
Kevin Gimpel, Dipanjan Das and Noah A. Smith
Recent speed-ups for training large-scale models like those found in statistical NLP exploit distributed computing (either on multicore or "cloud" architectures) and rapidly converging online learning algorithms. Here we aim to combine the two. We focus on distributed, "mini-batch" learners that make frequent updates asynchronously (Nedic et al., 2001; Langford et al., 2009). We generalize existing asynchronous algorithms and experiment extensively with structured prediction problems from NLP, including discriminative, unsupervised, and non-convex learning scenarios. Our results show asynchronous learning can provide substantial speedups compared to distributed and single-processor mini-batch algorithms with no signs of error arising from the approximate nature of the technique.

On Reverse Feature Engineering of Syntactic Tree Kernels
Daniele Pighin and Alessandro Moschitti
In this paper, we provide a theoretical framework for feature selection in tree kernel spaces based on gradient-vector components of kernel-based machines. We show that a huge number of fea-

tures can be discarded without a significant decrease in accuracy. Our selection algorithm is as accurate as and much more efficient than those proposed in previous work. Comparative experiments on three interesting and very diverse classification tasks, i.e. Question Classification, Relation Extraction and Semantic Role Labeling, support our theoretical findings and demonstrate the algorithm performance.

Inspecting the Structural Biases of Dependency Parsing Algorithms

Yoav Goldberg and Michael Elhadad

We propose the notion of a *structural bias* inherent in a parsing system with respect to the language it is aiming to parse. This structural bias characterizes the behaviour of a parsing system in terms of structures it tends to under- and over- produce. We propose a Boosting-based method for uncovering some of the structural bias inherent in parsing systems. We then apply our method to four English dependency parsers (an Arc-Eager and Arc-Standard transition-based parsers, and first- and second-order graph-based parsers). We show that all four parsers are biased with respect to the kind of annotation they are trained to parse. We present a detailed analysis of the biases that highlights specific differences and commonalities between the parsing systems, and improves our understanding of their strengths and weaknesses.

Index

– C –

– D –

– G –

– H –

– K –

– L –

– M –

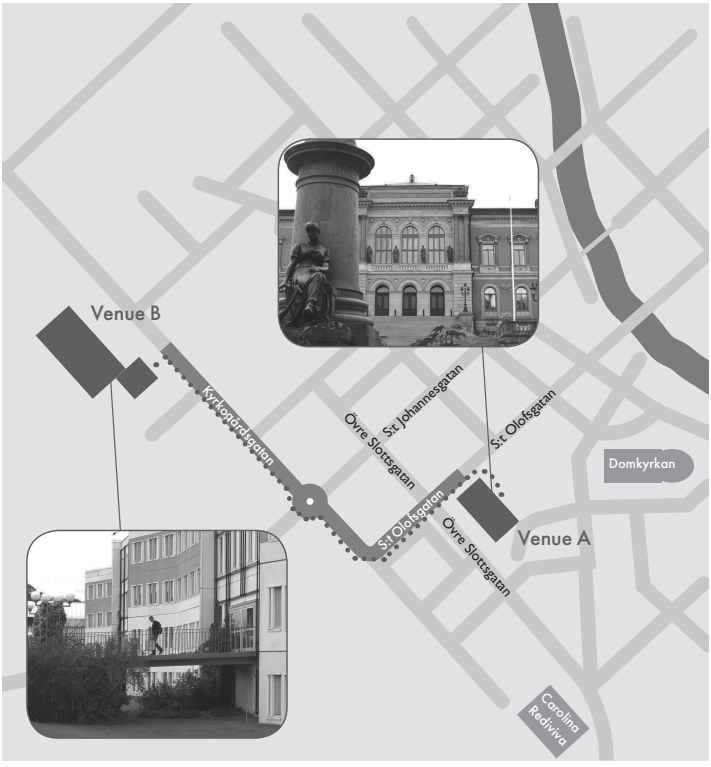
– Q –

– R –

– S –

– T –

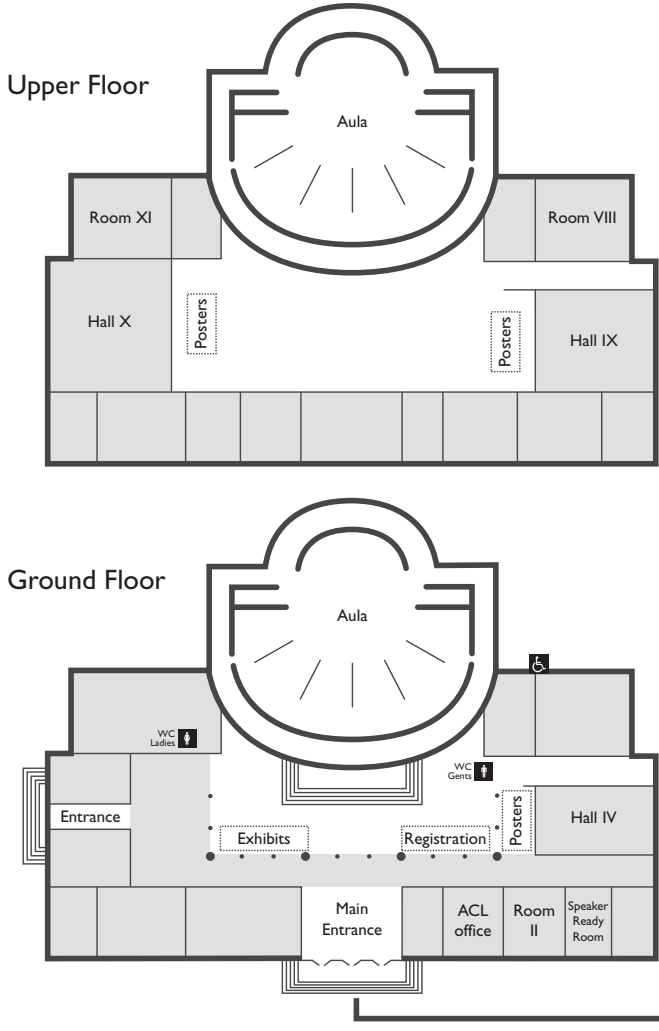
Venue Map

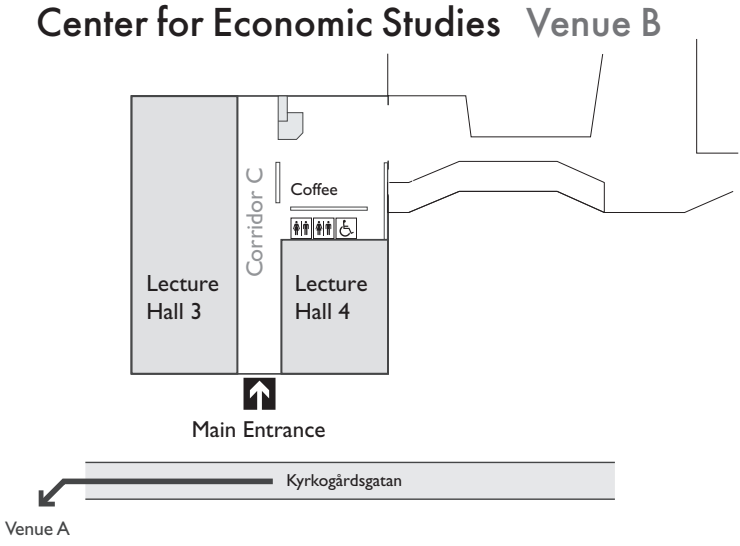


Uppsala University Main Building Venue A
Junction of Övre Slottsgatan and S:t Olofsgatan

Center for Economic Studies, Ekonomikum Venue B
Kyrkogårdsgatan 10

Uppsala University Main Building Venue A




Center for Economic Studies Venue B
Corridor C
Coffee
Lecture Hall 3
Lecture Hall 4
Main Entrance
Kyrkogårdsgatan
Venue A